

The price of imperfect competition for a spanning network

Hervé Moulin^a and Rodrigo A. Velez^b

^aDepartment of Economics, Rice University, Houston, TX 77251

^bDepartment of Economics, Texas A&M University, College Station, TX 77843

January 20, 2012

Abstract

A buyer procures a network to span a given set of nodes; each seller bids to supply certain edges, then the buyer purchases a minimal cost spanning tree. An efficient tree is constructed in any equilibrium of the Bertrand game.

We evaluate the *price of imperfect competition* (PIC), namely the ratio of the total cost that could be charged to the buyer in some equilibrium, to the true minimal cost. If each seller can only bid for a single edge and costs satisfy the triangle inequality, we show that the PIC is at most 2 for an odd number of nodes, and at most $2\frac{n-1}{n-2}$ for an even number n of nodes. Surprisingly, this worst case ratio does not improve when the cost pattern is ultrametric (a much more demanding substitutability requirement), although the overhead is much lower on average under typical probabilistic assumptions. But the PIC increases swiftly when sellers can only provide a subset of all edges.

Keywords: Algorithmic mechanism design; frugality; minimum cost spanning tree problem; price of imperfect competition.

Acknowledgments: Several comments by two anonymous referees, an associate editor, Tim Roughgarden, Jay Sethuraman, Vijay Vazirani, and seminar participants at CORE, CoED 2011, Louisiana State, Montreal, Rochester, and Texas A&M have been very helpful. Moulin's research is supported by the NSF under Grant CCF 1101202.

1 Introduction

Bertrand competition for differentiated commodities typically implies some welfare losses, as in the Hotelling model. We consider a special case where it does not, and where the surplus that the competing firms are able to extract admits a simple upper bound.

A buyer procures a network spanning a given set of nodes, while the sellers bid for different edges of the network. Efficiency requires to build a minimal cost

spanning tree. This familiar optimization problem has a variety of applications (including rail infrastructure, the internet’s backbone, water distribution, etc.; see [22] for a survey).

If one seller is the sole bidder for a certain edge, or a group of edges, he can typically bid higher than their true cost. But this overhead is bounded above because the edges are partial substitutes: for any edge e , there are several alternative paths ensuring the connection of e ’s two end-nodes. We wish to evaluate the welfare consequences of this *imperfect competition* between the sellers.

We observe first that irrespective of the ownership structure and of the cost pattern, in equilibrium an efficient (minimal cost) spanning tree is constructed (Proposition 1). This is very close to being a special case of the efficiency result in the matroid markets of [10] (on which more below).

Our main concern is to evaluate the surcharge to the buyer procuring the network. We provide tight bounds when the cost of the edges satisfy the triangular inequality, i.e., for any three edges e, f, g forming a triangle the cost of one edge is not larger than the sum of the costs of the other two. We speak then of *metric* costs. This familiar assumption is realistic when costs are a sub-additive function of the euclidean length of an edge, or some other measure of spatial distance between its end nodes (for instance the shortest length of a path connecting the nodes in an arbitrary network).¹ It is also automatic when all sellers have access to the same technology to build edges, but each one is only licensed to bid for certain edges: it is then feasible to connect the end nodes of e by building f and g .

Our first main result (Theorem 1) is that if each seller bids for a single edge, there is at least one bidder for each edge, and true costs are metric, the buyer’s surcharge essentially cannot exceed 100% of the true minimal cost. However our second main result (Theorem 2) says that the overhead grows rapidly if some edges are not available (no one is bidding for those edges): if at least half of the edges are missing, the worst case ratio of the overhead to the true cost is $n - 1$, where n is the number of nodes to connect.

It is instructive to compare our results to those of the literature evaluating the *frugality* of some related auction mechanisms ([2, 3, 17, 23]). The procurement of a spanning tree among sellers who each own an edge of the network is a special case of the procurement of a matroid basis; other examples include the procurement of a path between two given nodes of the network ([3, 11, 15]), and more generally the procurement of a team to perform a complex task [17]. Like us, these papers are poised to evaluate the worst possible surcharge to the buyer. However, the payments to the various sellers are more complicated than in the simple first-price auction of Bertrand competition: they are given by the canonical VCG mechanism (known as the *pivotal* mechanism since [13]), in which each winning edge is paid its true cost, plus the extra cost incurred if we cannot use that very edge in our spanning tree. Under this payment scheme bidding one’s true cost is a dominant strategy for each seller. Thus, implemen-

¹The cost of e takes the form $f(\|e\|)$, where f has increasing returns and $f(0) = 0$.

tation of the mechanism is *prior-free*, in sharp contrast with our equilibrium analysis of the Bertrand game requiring complete information among sellers.²

When we assume, as we do in our main results, that a seller bids for a single edge, it turns out that the two games, Bertrand bidding and pivotal, are essentially equivalent: the buyer’s surcharge in the pivotal game is precisely the same as in the most expensive equilibrium (equivalently, in its unique equilibrium in undominated strategies) of the Bertrand game (see Subsection 6.1). Thus, our results can be used indifferently in both contexts.

There have been at least three attempts in the literature to measure the frugality of a mechanism to procure a spanning tree. They differ in the choice for the benchmark cost to which the total payment of the seller is compared. In Archer and Tardos ([2, 3]), like here, the benchmark is the true minimal (efficient) cost. They show that the worst case ratio (total charge to efficient cost) is again $n - 1$, where n is the number of nodes to connect.³ Our Theorem 1 qualifies this negative result when costs are metric.

Subsequently, Talwar ([23]) uses for benchmark the cheapest cost of a spanning tree with no edge in common with the efficient tree. This curious proposal may lead in our model to a frugality ratio smaller than one!⁴ Finally in Karlin et al. ([17]) the benchmark is the solution of a linear program that, in the spanning tree problem, coincides with the most expensive equilibrium (for the buyer), so that the frugality ratio is one, tautologically.

We submit that from the buyer’s perspective, the only meaningful benchmark is the true efficient cost. To avoid confusion with the multiple frugality indices, and to convey the basic economic intuition, we propose to call *price of imperfect competition* (PIC) the ratio of the worst buyer’s charge in *some* equilibrium (or the actual charge in the pivotal mechanism), to the efficient cost.⁵

Summary of results

Section 2 introduces the minimum cost spanning tree problem and the Bertrand game of procurement. Proposition 1 states that an efficient tree is built in all relevant equilibria. Starting with section 3 we assume that each seller bids for only one edge. We provide a general formula for the PIC that uses only the network structure and arbitrary costs (Proposition 2).

We assume in section 4 that costs are metric, and show (Theorem 1) that if all edges have at least one bidder the worst case PIC is 2 for an odd number of nodes, and $2\frac{n-1}{n-2}$ for an even number n . However a single missing edge raises the worst case PIC to 3 (Proposition 3), and more generally the PIC reaches $n - 1$ as soon as half of the edges are missing (Theorem 2).

²As well as the technical twist of the limit equilibrium, see Section 3.

³Assume costs are infinite outside a single cycle connecting all n nodes, and on the cycle all costs are zero except for a single edge with cost 1. Then the ratio is $n - 1$.

⁴Assume 4 nodes a,b,d,e with metric costs $c_{ab} = c_{bd} = c_{de} = 1, c_{ad} = c_{be} = 2, c_{ae} = 3$. Efficient cost is 3, the total pivotal charge is 6, and the cost of the only edge disjoint tree is 7.

⁵This terminology is obviously inspired by the notions *price of anarchy* and *price of stability* ([16], [19]). But the PIC measures the impact of decentralizing the procurement process on the buyer’s welfare only. In our simple model imperfect competition induces no aggregate welfare loss.

In section 5 we assume that costs are *ultrametric*, i.e., if edges e, f, g form a triangle, $c_e \leq \max\{c_f, c_g\}$. This is stronger, and implies much closer substitutability than the triangular inequality. Ultrametric costs are a natural restriction in some environments. Consider a set of nodes for which the connection cost is determined by the compatibility of certain attributes that are ordered in decreasing levels of complexity. For instance the nodes are scattered in North America (consisting of 3 countries), and an edge has the largest cost θ_1 if the two endnodes are in two different countries, the second largest cost θ_2 if they are in the same country but different states or provinces, the third largest cost if in different counties of a given state or province, and so on.⁶

Ultrametric costs also play an important role in the design of rules sharing the cost of an optimal spanning tree between the different nodes (see [20], [8], [5], [6]).

Surprisingly, the worst case PIC does not improve (decrease) when the cost pattern is ultrametric, although it is clearly better on average. We propose some numerical computations to justify the latter.

In Section 6 we show first that if the familiar pivotal mechanism replaces our Bertrand game, it delivers the equilibrium of the latter most expensive for the buyer, therefore all our results apply. Next we explain how much of our results are preserved in the more general context of an “auction for a matroid basis” discussed above. Finally we probe the robustness of our results when a single seller can be the exclusive bidder on several edges. The news is good only if those edges are mutually disconnected.

Proofs are gathered in the Appendix.

2 Model

2.1 Minimum cost spanning trees: definitions and notation

Let $V \equiv \{1, \dots, n\}$ be a finite set of nodes with generic elements a, b, \dots , and G its associated set of edges, i.e., non oriented pairs in V , with generic elements e, f, \dots , and cardinality $\frac{n(n-1)}{2}$. A *network structure* on V is a pair $M \equiv (E, F)$ in which $E \subseteq G$ is a subset of edges that span V and $F \subseteq 2^E$ is the set of forests in E , i.e., subgraphs of E containing no cycle (we include F in the description of a network structure in order to make transparent the generalization of some of our results to “matroid” structures; see Section 6.2 for details). A circuit of M is a subset of E that is not in F and such that all of its proper subsets belong to F . Let $\Gamma_M \subseteq F$, with generic element γ , be the set of trees in E spanning V .

A cost “matrix” for $M \equiv (E, F)$, generically denoted by $c \in \mathbb{R}_+^E$, specifies a non-negative cost for each edge. For each $\gamma \in \Gamma_M$, let $c(\gamma) \equiv \sum_{e \in \gamma} c_e$ be the

⁶Formally each node in V is described by an ordered list of m attributes. The cost to connect two nodes is θ_k , where k is the first attribute in which they differ, and θ_k decreases in k . Such a cost matrix is ultrametric.

cost of γ at c ; the minimal cost of a tree in E spanning V is $\lambda(c) \equiv \min_{\gamma \in \Gamma_M} c(\gamma)$ and $\Gamma_M(c)$ is the set of minimal cost spanning trees in E for c . This set obtains by the well known Kruskal Algorithm (KA): (i) starting from the empty set, add one element in E that has minimum cost among the edges not yet chosen and whose addition creates a set still in F ; (ii) repeat as many times as possible [18].

Let $E(c)$ be the set of *efficient* edges at c , i.e., elements of E that belong to some tree in $\Gamma_M(c)$. We say that $e \in E(c)$ is *essential* if it belongs to each minimum cost spanning tree at c .

A cost matrix $c \in \mathbb{R}_+^E$ satisfies the triangle inequality on M if for each circuit of M , $T \subseteq E$, no edge $e \in T$ costs more than half the aggregate cost of T . If E is the complete graph G , the triangle inequality can be equivalently stated as: for any three edges forming a triangle, $\{e, f, g\} \subseteq E$, $c_e \leq c_g + c_f$. We speak of “metric costs for M ” and let $T(M)$ be the set of metric costs for M .

We use the following additional notation. The path between two nodes a, b on the spanning tree γ is denoted $[a, b]_\gamma$. Given $\gamma \in \Gamma_M$, $e \in \gamma$, and $f \in E$, we write $\gamma - e + f$ for the graph $(\gamma \setminus \{e\}) \cup \{f\}$. We write $\Delta(e, \gamma, M)$ for the set of edges $f \in E \setminus \gamma$ across e on γ , i.e., such that $f \neq e$ and $\gamma - e + f \in \Gamma_M$. We write $Ad(e, \gamma)$ for the set of edges adjacent to e in γ , i.e., $f \in \gamma \setminus \{e\}$ that share a node with e . For $c \in \mathbb{R}_+^E$, $b \in \mathbb{R}_+$, and $e \in E$, replacing the e -coordinate of c by b yields the cost matrix (c_{-e}, b) .

2.2 Preliminary results

We will use repeatedly several well known facts about minimum cost spanning trees. The first one holds for the more general matroid structures, where our model becomes the Bertrand competition for the procurement of a matroid basis: see Section 6.2 for details.

Lemma 1 *Let $M \equiv (E, F)$ be a network structure and $\{\gamma, \gamma'\} \subseteq \Gamma_M$. Then,*

1. ([9]; see Corollary 3) *There is a bijection $\varphi : \gamma \setminus \gamma' \rightarrow \gamma' \setminus \gamma$ such that for each $e \in \gamma \setminus \gamma'$, $\gamma - e + \varphi(e) \in \Gamma_M$.*
2. ([9]; Theorem 2) *For each $e \in \gamma$ there exists $f \in \gamma'$ such that $\{\gamma - e + f, \gamma' - f + e\} \subseteq \Gamma_M$.*

We observe next that the aggregate cost of an inefficient spanning tree can be reduced by replacing a single edge (recall that all proofs are gathered in the Appendix).

Lemma 2 *Let $M \equiv (E, F)$ be a network structure and $c \in \mathbb{R}_+^E$. Let $\gamma \in \Gamma_M \setminus \Gamma_M(c)$. Then, there is $e \in \gamma$ and $f \in E \setminus \gamma$ such that $c_f < c_e$ and $\gamma - e + f \in \Gamma_M$.*

Next, a minimum cost spanning tree preserves its optimality whenever the cost of one of its edges decreases; moreover, an efficient edge becomes essential if its cost decreases.

Lemma 3 *Let $M \equiv (E, F)$ be a network structure and $c \in \mathbb{R}_+^E$. Let $\gamma \in \Gamma_M(c)$, $e \in \gamma$, and $c'_e < c_e$. Then, $\gamma \in \Gamma_M(c_{-e}, c'_e)$ and e is essential for (c_{-e}, c'_e) .*

Finally, an efficient (resp. essential) edge remains so whenever the cost of the other elements in E increase.

Lemma 4 *Let $M \equiv (E, F)$ be a network structure and $c \in \mathbb{R}_+^E$.*

1. *Let $e \in E(c)$ and $c' \geq c$. Then, $e \in E(c'_{-e}, c_e)$.*
2. *Let $e \in E(c)$ be essential for c . Let $c' \geq c$. Then, e is essential for (c'_{-e}, c_e) .*

2.3 Bertrand competition

Let $M \equiv (E, F)$ be a network structure. We study the strategic situation in which a single buyer requests bids for each edge in E from a set of sellers S (with generic elements i, j, \dots) in order to construct a spanning tree of V . Each seller i is allowed (licensed) to bid for a subset of edges denoted $l(i)$. The sets $l(i)$ may overlap, and we assume $\cup_{i \in S} l(i) = E$. If only one seller is requested to bid for a given edge, we can think of him as the owner of the edge. We assume throughout that no seller has monopoly power: there is no seller who is the unique bidder for a set of edges that have non-empty intersection with each spanning tree of V , i.e., a set of edges containing a cut of M .

For each $e \in l(i)$, seller i can build e at a cost c_e^i . The efficient cost of edge e is $c_e^* \equiv \min_{\{i: e \in l(i)\}} c_e^i$.

A strategy of seller i is $p^i \equiv (p_e^i)_{e \in l(i)} \in \mathbb{R}_+^{l(i)}$, specifying his bid for each of “his” edges. Given a strategy profile $p \equiv (p^i)_{i \in S}$, the buyer purchases a cheapest spanning tree, namely $\gamma(p) \in \Gamma_M(p^*)$, where $p_e^* \equiv \min_{\{i: e \in l(i)\}} p_e^i$ is the lowest bid for edge e . The Bertrand game is well defined once we choose a tie-breaking rule for the case where $\Gamma_M(p^*)$ contains several cheapest trees, and for selecting a seller if there are several optimal bids for an edge $e \in \gamma(p)$. The net profit of seller i is then $\sum (p_e^i - c_e^i)$, where the sum bears over all edges e in $\gamma(p)$ for which i 's bid wins.

We keep the tie-breaking rule entirely general, and omit it for simplicity in the notation (it will be clear that our results are independent of this rule, that could even be probabilistic). Thus the Bertrand game is described by the triple $(M, S, (c^i)_{i \in S})$.

In this game with pure strategies, it is well known that the concept of Nash equilibria is not the right one: the payoff functions are discontinuous around a tie, implying that there may be no Nash equilibrium at all. The simplest example has two nodes and one edge, for which two sellers with different costs $c^1 < c^2$ compete (and ties broken in favor of seller 2, or randomly); see e.g. [24, Chapter 5].

This difficulty can be resolved in several convincing, equivalent ways. From the handful of possible approaches, the most convenient for our purpose turns out to be the notion of ε -equilibrium (see [12]), namely strategy profiles q from

which no seller can change his bid and increase his profit by more than ε .⁷ We call p a *limit equilibrium* if it is the limit of a sequence of ε -equilibria where ε goes to zero. In the single edge example, this implies identical bids $p^1 = p^2 \in [c^1, c^2]$ and that seller 1 wins the contract. We write $\mathcal{E}(M, S, (c^i)_{i \in S})$ for the set of limit equilibria of our game.

Note that in an ε -equilibrium q , the buyer selects a minimal cost tree $\gamma(q) \in \Gamma_M(q^*)$; thus, a sequence of ε -equilibria may have several limit points. Thus, in a given limit equilibrium p , several trees can be built, but of course they all cost $\lambda(p^*)$; similarly we could have more than one possible winner for each edge. These complications are totally inconsequential in the results below.

2.4 Efficiency of limit equilibria

Our first result states that in our model competition in prices ensures that an efficient network is built. Chen and Karlin ([10]) pioneered the strategic analysis of Bertrand competition in the procurement of a matroid basis, a more general problem than the procurement of a spanning tree, when each seller is the unique bidder for a single item. They showed that an efficient basis (a minimum cost spanning tree in our model) is built in each limit equilibrium of the Bertrand game in which no seller plays a weakly dominated strategy. The following proposition generalizes this result in two ways. First, efficiency holds for *all* limit equilibria. Second, it holds independently of the number of edges for which a seller can bid.

Proposition 1 *In each limit equilibrium of the Bertrand game $(M, S, (c_i)_{i \in S})$, the buyer purchases a minimal cost spanning tree $\gamma \in \Gamma_M(c^*)$ (though he typically pays more than $\lambda(c^*)$).*

In Section 6.2 we generalize Proposition 1 to the procurement of a matroid basis.

2.5 The price of imperfect competition

We define the Price of Imperfect Competition (PIC) as the ratio of the maximal aggregate price paid by the buyer in a limit equilibrium of the Bertrand game $(M, S, (c^i)_{i \in S})$, to the true minimal cost $\lambda(c^*)$.⁸

$$PIC(M, S, (c^i)_{i \in S}) \equiv \frac{\max \{ \lambda(p^*) : p \in \mathcal{E}(M, S, (c^i)_{i \in S}) \}}{\lambda(c^*)},$$

(where we use max instead of sup because we show below that the maximum is always reached).

⁷Other solutions include: to extend the definition of equilibria and make the tie breaker an outcome of agents' strategic interaction ([21]); to eliminate weakly dominated strategies ([7]); to assume integer-valued bids; to use mixed strategies ([14], Chapter 10).

⁸We adopt the convention $0/0 = 1$ and for each $b > 0$, $b/0 = +\infty$.

3 Computing the PIC

In this section we characterize limit equilibria of the Bertrand game under the assumption that each seller bids for exactly one edge. This implies a general expression of the PIC, the key to our subsequent computation of the PIC under metric or ultrametric costs.

Fix a network structure M , a cost matrix c , and an efficient edge $e \in E(c)$. For our purpose a key quantity is how high the cost of e can be, ceteris paribus, while e remains in at least one minimal cost spanning tree:

$$\mu_e(c) \equiv \max \{b \in \mathbb{R}_+ : e \in E(c_{-e}, b)\}. \quad (1)$$

Lemma 5 *Let $M \equiv (E, F)$ be a network structure, $c \in \mathbb{R}_+^E$, and $e \in E(c)$. Then,*

(i) *The cost $\mu_e(c)$ is the minimum cost among the substitutes of e : for each $\gamma \in \Gamma_M(c)$ such that $e \in \gamma$, $\mu_e(c) = \min\{c_f : f \in \Delta(e, \gamma, M)\}$.*

(ii) *Essential edges are the only ones whose cost can be increased and still remain efficient: if e is essential, then $\mu_e(c) > c_e$; if e is not essential, then $\mu_e(c) = c_e$.*

(iii) *The sum of the costs $\mu_e(c)$ along a minimal cost spanning tree is an invariant of the cost matrix c . For all $\gamma, \gamma' \in \Gamma_M(c)$ we have*

$$\sum_{e \in \gamma} \mu_e(c) = \sum_{e' \in \gamma'} \mu_{e'}(c),$$

so this sum can be written as $\mu(c)$. Moreover, $\mu(c)$ is weakly increasing in c .

We can now describe the range of the aggregate prices that the buyer may be charged in limit equilibria of the Bertrand game. We highlight the case where each seller is the exclusive bidder for exactly one edge, because for a given cost matrix, it happens to deliver the highest possible charge to the buyer. We speak then of the *exclusive case*, and identify each seller with his edge, so that S drops from the notation and we simply write $PIC(M, c)$.

Proposition 2 *Consider the Bertrand game $(M, S, (c^i)_{i \in N})$.*

(i) *Suppose that each seller bids for exactly one edge. Then in each limit equilibrium of $(M, S, (c^i)_{i \in N})$, the buyer pays at least $\lambda(c^*)$ and at most $\mu(c^*)$ (recall that $c_e^* = \min_{\{i: e=l(i)\}} c_e^i$).*

(ii) *Suppose that each seller is the exclusive bidder for exactly one edge, so we identify a seller with his edge e and c_e is this bidder's cost. Then for each $\gamma \in \Gamma_M(c)$, the profile*

$$p_e = \mu_e(c) \text{ if } e \in \gamma; p_e = c_e \text{ if } e \notin \gamma, \quad (2)$$

is a limit equilibrium of $(M, S, (c^i)_{i \in N})$ where the buyer pays $\mu(c)$.

Note that if in the exclusive case we refine the Nash equilibrium concept by one round of elimination of weakly dominated strategies, then in all remaining

limit equilibria the buyer pays exactly $\mu(c)$. This is because bidding below one's cost c_e^i is a weakly dominated strategy for bidder i , so that all limit equilibria are described by (2), where γ varies in $\Gamma_M(c)$.

Corollary 1 *Consider the Bertrand game $(M, S, (c^i)_{i \in N})$.*

i) Suppose that each seller bids for exactly one edge. Then,

$$PIC(M, S, (c^i)_{i \in N}) \leq \frac{\mu(c^*)}{\lambda(c^*)}.$$

ii) Suppose that each seller is the exclusive bidder for exactly one edge. Then,

$$PIC(M, c) = \frac{\mu(c)}{\lambda(c)}.$$

As a first application, the Corollary explains the knife-edge cost structures where the buyer pays no overhead at all, i.e., $PIC(M, S, (c^i)_{i \in N}) = 1$. This holds if $\lambda(c^*) = \mu(c^*)$. By Lemmas 5 (ii) and 5 (iii), we can formulate this condition in more transparent ways.

Lemma 6 *The overhead is zero in all limit equilibria of $(M, S, (c^i)_{i \in S})$ if, and only if in the exclusive case, one of the three following equivalent conditions holds:*

i) $\lambda(c^) = \mu(c^*)$;*

ii) every edge in $E(c^)$ is non essential;*

iii) at each stage of each instance of KA for M at c , we can select at least two edges.

4 Metric costs

The overhead that the buyer can expect to pay above the true minimal cost depends on the degree of substitutability of the edges. For instance, suppose that for some cost matrix c there is a unique minimal cost spanning tree γ . If costs outside γ are unboundedly large, then there is no substitute for any edge in γ . Hence the overhead payed by the buyer is infinite. At the other extreme, if *all* edges in G have the same cost, Lemma 5 (i) gives $\mu(c) = \lambda(c)$ so the overhead is zero (despite the fact that edges are *not* perfect substitutes).

The triangle inequality captures a more realistic form of substitutability: the cost of each edge is bounded by the aggregate cost of the paths that substitute it. The three-node case illustrates that under this intermediate notion of substitutability, the overhead that the buyer may end up paying is a “small” factor of the true cost. Indeed, suppose that $G \equiv \{e, f, g\}$ and that $\{e, f\}$ form a minimum cost spanning tree at c . If costs are metric, then $c_g \leq c_e + c_f$. Thus, both $\mu_e(c)$ and $\mu_f(c)$ are at most $c_e + c_f$, and $\mu(c) \leq 2\lambda(c)$, i.e., the PIC is at most 2.

Our main result is that this is essentially the worst possible PIC over all metric costs when each seller bids for just one edge and all edges are offered by at least one seller.

We start with a simple upper bound independent of the network structure.

Proposition 3 *If in the Bertrand game $(M, S, (c^i)_{i \in S})$ the cost matrix c^* is metric, i.e., $c^* \in T(M)$, then*

$$PIC(M, S, (c^i)_{i \in S}) \leq n - 1.$$

Our main results, Theorems 1 and 2, show that this bound is reached for sparse graphs E , specifically those with less than half the number of edges in the complete graph, whereas for the complete graph the bound is essentially 2.

4.1 Complete graph

Theorem 1 *Let $M \equiv (G, F)$ be the complete network structure. Assume that each seller is the exclusive bidder for exactly one edge and the cost matrix c is metric, i.e., $c \in T(M)$. Then,*

$$\max_{c \in T(M)} PIC(M, c) = \begin{cases} 2 & \text{if } n \text{ is odd,} \\ 2\frac{n-1}{n-2} & \text{if } n \text{ is even.} \end{cases}$$

Recall from Corollary 1 that this upper bound applies as well to any non exclusive game $(M, S, (c^i)_{i \in S})$ provided c^* is metric.⁹ We state all subsequent results in this section and the next in the exclusive case, keeping in mind that they apply to the non exclusive case as well.

We give now examples where the two bounds in Theorem 1 are reached, and defer to the appendix the proof that they are indeed upper bounds.

Example 1 *Let $M \equiv (G, F)$ be the complete network structure and $\gamma \in \Gamma_M$ an arbitrary spanning tree. Consider the cost matrix defined by: for each $e \in G$, $c_e \equiv 1$ if $e \in \gamma$, and $c_e \equiv 2$ otherwise. Clearly, c is metric. By Lemma 5 (i), for each $e \in \gamma$, $\mu_e(c) = 2$. Hence, $\mu(c) = 2(n - 1)$, while $\lambda(c) = (n - 1)$. By Corollary 1, $PIC(M, c) = 2$.*

Example 1 shows that for each spanning tree γ , there is a cost matrix c such that $\gamma \in \Gamma_M(c)$ and $PIC(c) = 2$. Interestingly, for n even, choosing c such that $PIC(c) = 2\frac{n-1}{n-2}$ places many more constraints on the structure of the optimal tree. We start with an example.

Example 2 *Let $M \equiv (G, F)$ be the complete network structure. Let γ be the line-tree with edges $i(i + 1)$ for $1 \leq i \leq n - 1$. Set $c_{i(i+1)} \equiv 0$ if i is odd, $c_{i(i+1)} \equiv 1$ if i is even, so we have $\frac{n}{2}$ zero cost edges, including the two end edges. Set $c_{i(i+2)} \equiv 1$ for $1 \leq i \leq n - 2$, and $c_e \geq 1$ for other edges (we do not need to specify their costs as long as they remain metric). By Lemma 5 (i), $\mu_e(c) = 1$ for all $e \in \gamma$. Hence $\mu(c) = n - 1$, while $\lambda(c) = \frac{n}{2} - 1$. By Corollary 1, $PIC(M, c) = 2\frac{n-1}{n-2}$.*

⁹The bound is reached at a cost profile where for each edge all but one seller have very large costs.

Example 2 shows that when n is even, given any linear tree, there exists metric costs reaching the bound in Theorem 1 and for which the tree is optimal. The following proposition characterizes all trees sharing this property.

Proposition 4 *Let $M \equiv (G, F)$ be the complete network structure. Assume that each seller is the exclusive bidder for exactly one edge.*

i) If n is odd, every spanning tree $\gamma \in \Gamma_M$ admits a cost matrix $c \in T(V)$ such that $\gamma \in \Gamma_M(c)$ and $PIC(M, c) = 2$.

ii) If n is even, the spanning tree $\gamma \in \Gamma_M$ admits a cost matrix $c \in T(V)$ such that $\gamma \in \Gamma_M(c)$ and $PIC(M, c) = 2\frac{n-1}{n-2}$ if and only if the edges of γ contain a perfect matching of V (i.e., we can find $\frac{n}{2}$ edges in γ such that each node in V is the end point of exactly one such edge).

Remark 1 *An easy variant of Theorem 1 obtains if we relax the triangle inequality by a factor $\rho \geq 1$: for any three edges forming a triangle, $c_e \leq \rho(c_f + c_g)$. Then the upper bounds 2 and $2\frac{n-1}{n-2}$ are simply multiplied by ρ . The straightforward proof mimics that of Theorem 1.*

4.2 Incomplete graph

In this section we investigate the PIC under metric costs when some edges may not be offered by the sellers. We begin with the case in which only one edge is missing, for which we compute the worst case PIC exactly.

The following example has $n \geq 6$, one edge is not available for purchase, and metric costs for which the PIC is 3.

Example 3 *(Figure 1 (a)) Assume that $n = 4$ or $n \geq 6$. Let $M \equiv (E, F)$ where $E \equiv G \setminus \{13\}$, i.e., E contains all edges except the edge that connects nodes 1 and 3. Let γ be the line-tree with edges $i(i+1)$ for $1 \leq i \leq n-1$. Let $c_{34} \equiv 1$ and for each $i \neq 3$, let $c_{i(i+1)} \equiv 0$. For each pair $ij \in E \setminus \gamma$, let $c_{ij} \equiv \sum_{e \in [i,j]_\gamma} c_e$. Clearly, c is metric. By Lemma 5 (i), $\mu_{(i,i+1)}(c) = 1$ for $i = 1, 2, 3$ and $\mu_{(i,i+1)}(c) = 0$ for $i > 3$. Hence $\mu(c) = 3$, while $\lambda(c) = 1$. By Corollary 1, $PIC(M, c) = 3$.*

The following theorem states that in fact this is the worst case when $n \geq 6$.

Proposition 5 *Assume that $n = 4$ or $n \geq 6$ and that each seller is the exclusive bidder for exactly one edge. Then,*

$$\max_{c \in T(M), M \equiv (E, F), |E|=|G|-1} PIC(M, c) = 3.$$

Proposition 3 shows the worst PIC when $n = 5$ is at most 4. The construction in Example 3 gives a metric cost for which $PIC(M, c) = 4$.

We now generalize the construction in Example 3 to the case in which more edges are missing. Intuitively, we maximize the number of edges e whose real cost is zero and for which $\mu_e(c) = 1$, while maintaining $\lambda(c) = 1$.

Example 4 Assume that $n \geq 6$. Let $k \in \{1, \dots, \lfloor \frac{n}{2} \rfloor - 2\}$ ($\lfloor \frac{n}{2} \rfloor$ is the greatest integer bounded above by $\frac{n}{2}$). Let $E_1 \equiv \{ij : i = 1, \dots, k, j = i + 2, \dots, k + 2\}$ and $M \equiv (E, F)$ where $E \equiv G \setminus E_1$. Let γ be the line-tree with edges $i(i + 1)$ for $1 \leq i \leq n - 1$. Let $c_{(k+2)(k+3)} \equiv 1$ and for each $i \neq k + 2$, let $c_{i(i+1)} \equiv 0$. For each pair $ij \in E \setminus \gamma$, let $c_{ij} \equiv \sum_{e \in [i,j]_\gamma} c_e$. Clearly, c is metric. By Lemma 5 (i), $\mu_{i(i+1)}(c) = 1$ for $i = 1, \dots, k + 2$ and $\mu_{ij}(c) = 0$ for $i = k + 3, \dots, n - 1$. Hence $\mu(c) = 2 + k$, while $\lambda(c) = 1$. By Corollary 1, $PIC(M, c) = 2 + k$.

Finally, the number of available edges is $|E| = \frac{n(n-1)}{2} - \frac{k(k+1)}{2}$. Therefore for n large, $PIC(M, c) \simeq 2 + \sqrt{2(\frac{n(n-1)}{2} - |E|)}$ where E contains at least $\frac{3}{4}$ of the edges.

Example 5 Assume that $n > 6$. Let $k \in \{1, \dots, n - (\lfloor \frac{n}{2} \rfloor + 3)\}$. Let $E_1 \equiv \{ij : i = 1, \dots, \lfloor \frac{n}{2} \rfloor - 2, j = i + 2, \dots, \lfloor \frac{n}{2} \rfloor\}$ and $E_2 \equiv \{(\lfloor \frac{n}{2} \rfloor + i)j : i = 1, \dots, k, j = \lfloor \frac{n}{2} \rfloor + i + 2, \dots, n\}$. Let $M \equiv (E, F)$ where $E \equiv G \setminus (E_1 \cup E_2)$. Let γ be the line-tree with edges $i(i + 1)$ for $1 \leq i \leq n - 1$. Let $c_{\lfloor \frac{n}{2} \rfloor(\lfloor \frac{n}{2} \rfloor + 1)} \equiv 1$ and for each $i \neq \lfloor \frac{n}{2} \rfloor$, let $c_{i(i+1)} \equiv 0$. For each pair $ij \in E \setminus \gamma$, let $c_{ij} \equiv \sum_{e \in [i,j]_\gamma} c_e$. Clearly, c is metric. By Lemma 5 (i), $\mu_{i(i+1)}(c) = 1$ for $i = 1, \dots, \lfloor \frac{n}{2} \rfloor + k$ and $\mu_{i(i+1)}(c) = 0$ for $i = \lfloor \frac{n}{2} \rfloor + k + 1, \dots, n - 1$. Hence $\mu(c) = \lfloor \frac{n}{2} \rfloor + k$, while $\lambda(c) = 1$. By Corollary 1, $PIC(M, c) = \lfloor \frac{n}{2} \rfloor + k$.

Finally, $|E| = |G| - \frac{1}{2}(\lfloor \frac{n}{2} \rfloor - 2)(\lfloor \frac{n}{2} \rfloor - 1) - k(n - \lfloor \frac{n}{2} \rfloor - \frac{k+3}{2})$; for instance if n is even $|E| = \frac{1}{8}(3n^2 + 2n - 8) - \frac{1}{2}k(n - k - 3)$. This time E contains between $\frac{3}{4}$ and $\frac{1}{2}$ of the edges, and for n large $PIC(M, c) \simeq n - \sqrt{2|E| - \frac{n(n-1)}{2}}$.

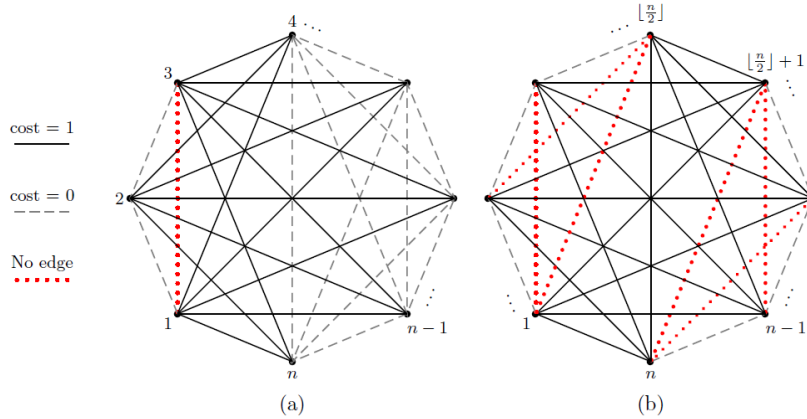


Figure 1: $M \equiv (E, F)$; (a) $PIC(c) = 3$ for $|E| = |G| - 1$; (b) $PIC(c) = n - 1$ for $|E| = \frac{1}{4}(n^2 + 4n - 8 - \varepsilon)$ where ε is 0 for even n and 1 for odd n .

Example 6 (Figure 1 (b)) Assume that $n \geq 6$. Let $k \equiv n - (\lfloor \frac{n}{2} \rfloor + 2)$. Let $E_1 \equiv \{ij : i = 1, \dots, \lfloor \frac{n}{2} \rfloor - 2, j = i + 2, \dots, \lfloor \frac{n}{2} \rfloor\}$. Let $E_2 \equiv \{(\lfloor \frac{n}{2} \rfloor + i)j : i = 1, \dots, k, j = \lfloor \frac{n}{2} \rfloor + i + 2, \dots, n\}$. Let $M \equiv (E, F)$ where $E \equiv G \setminus (E_1 \cup E_2)$. Let γ be the line-tree with edges $i(i + 1)$ for $1 \leq i \leq n - 1$. Let $c_{\lfloor \frac{n}{2} \rfloor(\lfloor \frac{n}{2} \rfloor + 1)} \equiv 1$ and

for each $i \neq \lfloor \frac{n}{2} \rfloor$, let $c_{i(i+1)} \equiv 0$. For each pair $ij \in E \setminus \gamma$, let $c_{ij} \equiv \sum_{e \in [i,j]_\gamma} c_e$. Clearly, c is metric. By Lemma 5 (i), $\mu_{i(i+1)}(c) = 1$ for $i = 1, \dots, n-1$. Hence $\mu(c) = n-1$, while $\lambda(c) = 1$. By Corollary 1, $PIC(M, c) = n-1$.

Finally, we compute $|E| = \frac{1}{4}(n^2 + 4n) - 8 - \varepsilon$, where ε is 0 for even n and 1 for odd n . Thus we reach the worst possible bound $n-1$ when E is missing less than half of the edges.

The following proposition summarizes the lower bound for the worst case PIC provided by Examples 3-6. We conjecture that this proposition also provides the upper bound for the PIC. Figure 2 illustrates it.

We write $\mathcal{N}(m)$ for the set of network structures $M \equiv (E, F)$ such that $|E| = m$, and $\overline{PIC}_m \equiv \max_{M \in \mathcal{N}(m), c \in T(M)} PIC(M, c)$.

Theorem 2 *Let $m \in \{1, \dots, |G| - 1\}$, and assume that each seller is the exclusive bidder for exactly one edge. Then, \overline{PIC}_m is bounded below as follows: for $1 \leq k \leq \lfloor \frac{n}{2} \rfloor - 2$ we have,*

$$\overline{PIC}_m \geq \begin{cases} 2+k & \text{for } m \leq \frac{n(n-1)}{2} - \frac{k(k+1)}{2} \\ \lfloor \frac{n}{2} \rfloor + k & \text{for } m \leq \frac{1}{8}(3n^2 + (1+\varepsilon)(2n-7) - 1) - \frac{1}{2}k(n-k-3+\varepsilon). \end{cases}$$

Moreover,

$$\overline{PIC}_m = n-1 \quad \text{for } m \leq \frac{1}{4}(n^2 + 4n) - 8 - \varepsilon,$$

(recall the notation $\varepsilon = n$ modulo 2).

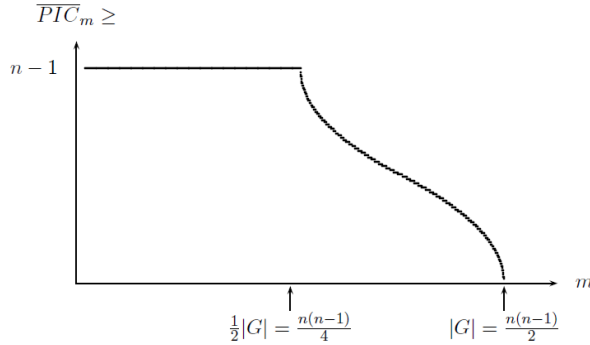


Figure 2: Lower bound for \overline{PIC}_m (values calculated for $n = 100$).

5 Ultrametric costs

We turn to a subdomain of metric costs that exhibit a higher degree of substitutability than the metric domain. We restrict attention to the benchmark case in which all edges are available for purchase by the buyer.

Let $M \equiv (G, F)$ be the complete network structure and $c \in \mathbb{R}_+^G$. We say that c is *ultrametric* if $c_e \leq \max\{c_f, c_g\}$ for any three edges e, f, g forming a triangle. The set $U(M)$ of ultrametric costs for M is a subset of $T(M)$.

Ultrametric costs emerge naturally in the minimum cost spanning tree model (see also the discussion of this assumption in the introduction). Fix any spanning tree $\gamma \in \Gamma_M$ and a profile of costs $c_\gamma \in \mathbb{R}_+^\gamma$ for the edges of γ only. From the point of view of the sellers bidding on those edges, the worst competition from sellers bidding on other edges is when their costs $\tilde{c}_{G \setminus \gamma}$ are the smallest possible such that γ is still in $\Gamma_M(c_\gamma, \tilde{c}_{G \setminus \gamma})$. It is well known (e.g., [6]) that this worst case cost $\tilde{c}_{G \setminus \gamma}$ is well defined, and is also the unique cost on $G \setminus \gamma$ such that $(c_\gamma, \tilde{c}_{G \setminus \gamma})$ is ultrametric and $\gamma \in \Gamma_M(c_\gamma, \tilde{c}_{G \setminus \gamma})$.

Intuitively, the higher level of substitutability of an ultrametric cost should imply a lower worst case PIC than a metric cost. Indeed, given $u \in U(M)$ and $\gamma \in \Gamma_M(u)$, there is a largest $c \in T(M)$, generally greater than u , such that for each $e \in \gamma$, $c_e = u_e$. It is easy to construct examples in which $\mu(u)$ is considerably less than $\mu(c)$.

Example 7 Consider a line-tree γ with nodes $1, \dots, n$, and $c_{12} < c_{23} < \dots < c_{(n-1)n}$. For each $ij \in G$, $i < j$, let $\tilde{c}_{ij} = c_{(j-1)j}$ (this is the cost \tilde{c} discussed above). For edge 12 , the overhead is $c_{23} - c_{12}$. For $i \geq 2$, the cheapest edge in $Ad(i(i+1), \gamma)$ is $(i-1)i$, implying that the overhead on $i(i+1)$ is zero. Therefore, $\mu(\tilde{c}) - \lambda(\tilde{c}) = c_{23} - c_{12}$. By contrast, let \bar{c} be the largest metric cost compatible with the given costs on γ , i.e., $\bar{c}_{ij} \equiv \sum_{k=i}^{j-1} c_{k(k+1)}$ for all $i < j$. It is easy to compute $\mu(\bar{c}) - \lambda(\bar{c}) = c_{12} + 2c_{23} + c_{34} + \dots + c_{(n-2)(n-1)}$. For instance if $c_{i(i+1)} = i$ for $1 \leq i \leq n-1$, we find $PIC(M, \tilde{c}) = 1 + \frac{2}{n(n-1)}$ while $PIC(M, \bar{c}) \geq 1.8$ for all n .

However, the *worst case PIC* is the same for metric and ultrametric costs. Since $U(V) \subseteq T(V)$, the claim follows if we show that the bounds in Theorem 1 can be reached in $U(V)$. Consider the line tree γ with alternating costs 0 and 1, and 0 at both end edges if n is even, and set a cost of 1 for any edge outside γ . This cost c is ultrametric. One can easily see that $PIC(M, c) = 2$ for odd n and $PIC(M, c) = 2\frac{n-1}{n-2}$ for n even.

There are some structural differences between the metric and ultrametric domains. The following analog of Proposition 4 states necessary and sufficient conditions for a tree to be optimal at some ultrametric costs reaching the worst case PIC.

Proposition 6 Let $M \equiv (G, F)$ be the complete network structure. Assume that each seller is the exclusive bidder for exactly one edge.

i) Let $\gamma \in \Gamma_M$. Then, γ admits a cost matrix $c \in U(V)$ such that $\gamma \in \Gamma_M(c)$ and $PIC(M, c) = 2$ if and only if γ has at least one leaf edge of which the inner end point is of degree two.

ii) If n is even, the spanning tree $\gamma \in \Gamma_M$ admits a cost matrix $c \in U(V)$ such that $\gamma \in \Gamma_M(c)$ and $PIC(M, c) = 2\frac{n-1}{n-2}$ if and only if the edges of γ contain a perfect matching of V .

Thus we can reach a PIC of 2 with a metric cost for any tree, whereas for a tree not covered by statement *i*) above, the PIC one can reach with an ultrametric cost is smaller, and sometimes much smaller. The most extreme example is the simple star, where all nodes but one are of degree one: the worst PIC of an ultrametric cost for which this star is optimal is $1 + \frac{1}{n-2}$. It is achieved when the cost of all edges is one, with the single exception of one edge adjacent to the center with cost zero.¹⁰ This observation suggests that the “expected” PIC is lower in the ultrametric domain than in the metric domain. We confirm this intuition by calculating the expected PIC in a simple probabilistic model.

To each cost matrix $c \in \mathbb{R}_+^G$ we associate the ultrametric cost $\tilde{c} \in U(V)$ defined as the the smallest cost bounded above by c and such that $\lambda(c) = \lambda(\tilde{c})$.¹¹ Let $C \subseteq T(V)$ be a set of triangular matrices (endowed with an appropriate σ -algebra) and p a probability measure on C . We associate with (C, p) the projected probability space (U, q) by the mapping $c \rightarrow \tilde{c}$. The two spaces, (C, p) and (U, q) are comparable in the sense that for each cost matrix $c \in C$ there is a cost matrix $\tilde{c} \in U$ with equivalent cost structure, i.e., $\Gamma_M(c) \subseteq \Gamma_M(\tilde{c})$ and $\lambda(c) = \lambda(\tilde{c})$.

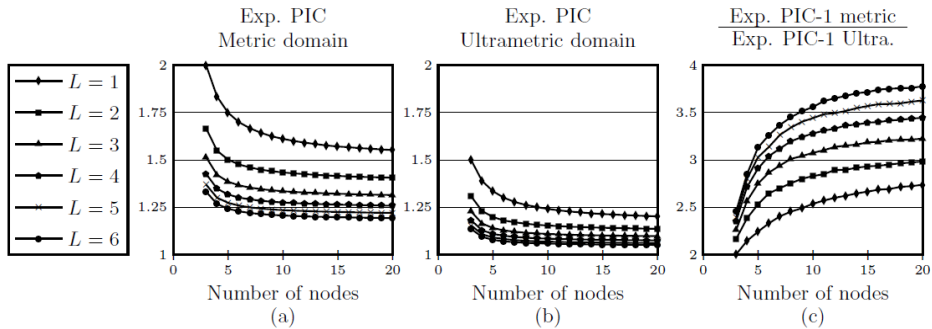


Figure 3: Expected PIC and expected overhead unit cube model.

We numerically calculate the expected PIC, for the triangular spaces and their associated ultrametric spaces, when the nodes $\{1, \dots, n\}$ are i.i.d uniformly in the unit cube $[0, 1]^L$, and costs correspond to the euclidean distance. Formally

$$C \equiv \{c \in \mathbb{R}^E : \text{for each } e = ab \in E, c_e = \|a - b\|, a, b \in [0, 1]^L\}, \quad (3)$$

and p is the Lebesgue measure. In contrast to the maximum PIC (which is essentially constant as a function of the number of nodes), the expected PIC, both in the triangular and ultrametric spaces, is decreasing with respect to the number of nodes. Figure 3 (a-b) shows the expected PIC in the two domains for $L \in \{1, \dots, 6\}$ and $n \in \{3, \dots, 20\}$.

¹⁰The proof of statement (ii) in Proposition 6 shows that this is the worst case for the star shape.

¹¹If we pick any optimal tree $\gamma \in \Gamma_M(c)$, the cost \tilde{c} is the same we discussed four paragraphs before Proposition 6. See [6].

We see next on Figure 3 (c) that the ratio of the expected overhead $PIC(M, c)$ —1 in the triangular domain, to that in the ultrametric domain, is increasing in n with an initial value of at least 2 for $n = 3$. Thus in our probabilistic model the expected overhead for metric costs is at least twice that for ultrametric costs.

6 Discussion

6.1 The pivotal interpretation of the PIC

We describe the pivotal mechanism using the assumptions and notation of Section 2.3. As before, the strategy of seller i is $p^i \equiv (p_e^i)_{e \in l(i)} \in \mathbb{R}_+^{l(i)}$, specifying his bid for each of “his” edges. Given a strategy profile $p \equiv (p^i)_{i \in S}$, the buyer purchases a cheapest spanning tree $\gamma(p) \in \Gamma_M(p^*)$, where p^* is the lower contour of bids (recall $p_e^* \equiv \min_{\{i: e \in l(i)\}} p_e^i$). We write $p^*(-i)$ for the lower contour of bids after we raise all of seller i 's bids to ∞ (note that $p_e^*(-i) = \infty$ happens for some $e \in E(p^*(-i))$ only if i is the sole bidder for a cut of the network). Then the buyer pays to seller i

$$\lambda(p^*(-i)) - \lambda(p^*) + \sum p_e^*$$

where the sum bears on all edges in $l(i) \cap \gamma(p)$ that i wins. Note that the additional payment $\lambda(p^*(-i)) - \lambda(p^*)$ is zero if and only if for each $e \in l(i)$ such that $p_e^i = p_e^*$, edge e is not essential at p^* and/or there is another winning bid for e .

It is well known that in this mechanism the truthful bid $(c_e^i)_{e \in l(i)}$ is a dominant strategy for each i , and the choice of an optimal tree does not matter utility-wise ([13]). At the true profile of costs the winning bid for edge e is $c_e^* \equiv \min_{\{i: e \in l(i)\}} c_e^i$ and the buyer purchases an efficient tree. His total payment is

$$\lambda(c^*) + \sum_N (\lambda(c^*(-i)) - \lambda(c^*)), \quad (4)$$

where $c^*(-i)$ is the lower contour of the costs after we raise all of seller i 's costs to ∞ .

We state now the counterpart of Proposition 2 for the pivotal mechanism.

Proposition 7 *i) Suppose that each seller bids for exactly one edge. Then for any $(c_{l(i)}^i)_{i \in S} \in \mathbb{R}_+^S$, in the pivotal mechanism the buyer pays at least $\lambda(c^*)$ and at most $\mu(c^*)$.*

ii) Suppose that each seller is the exclusive bidder for exactly one edge, so we identify a player with his edge e . Then for any $c \in \mathbb{R}_+^E$ the buyer pays $\mu(c)$ in the pivotal mechanism.

6.2 Generalization to Minimum cost matroid basis problem

The minimum cost spanning tree problem is a particular case of a more general optimization problem that we describe next. For simplicity we maintain the

notation developed for minimum cost spanning tree problems. The applications covered by these more general structures, known as “matroids,” also include bidding for a path connecting two given nodes in a network, hiring a team of agents, multi-units auctions, and more: papers [4] and [10] are among the first to discuss the performance of allocation mechanisms in matroid environments.

We explain now which of our results extend to arbitrary matroid structures. A matroid is a pair $M \equiv (E, F)$ in which E is a finite set and F is a family of subsets of E such that: (i) $\emptyset \in F$; (ii) each subset of a set in F is in F ; and (iii) for each pair of elements of F , A and B , such that $|A| < |B|$, there is $b \in B \setminus A$ such that $A \cup \{b\} \in F$.¹² A basis of M is a maximal set in F , i.e., an element of F that is not a proper subset of another element in F . All bases of M have the same cardinality, and this number is the dimension of M . Let Γ_M be the set of bases of M .

A cost “matrix” on M assigns a non-negative number to each element of E . The generic cost matrix is $c \in \mathbb{R}_+^E$. For each $\gamma \in \Gamma_M$ let $c(\gamma) \equiv \sum_{e \in \gamma} c_e$ be the cost of γ at c ; the minimal cost of a basis of M is $\lambda(c) \equiv \min_{\gamma \in \Gamma_M} c(\gamma)$ and $\Gamma_M(c)$ is the set of minimal cost bases of M for c .¹³

KA can be applied to a pair $M \equiv (E, F)$ in which E is a finite set and F is a family of subsets of E as follows: (i) starting from the empty set, add one element of E that has minimum cost among the elements not yet chosen and whose addition creates a set in F ; (ii) repeat as many times as possible. It is well known that for each matroid $M \equiv (E, F)$ and each $c \in \mathbb{R}_+^E$, KA obtains $\Gamma_M(c)$.

Bertrand competition for the procurement of a matroid basis by a buyer from a set of sellers S can be modeled in the same fashion as in the procurement of a spanning tree in a network structure (the no monopoly condition requires here that no agent be the unique bidder for a subset of E that has non-empty intersection with each basis of M). Here a Bertrand game is also described by the list $(M, S, (c^i)_{i \in S})$ in which for each $i \in S$, $c^i \in \mathbb{R}_+^{l(i)}$ represents the costs for the elements in E that seller i can build. Likewise, limit equilibrium is the right equilibrium concept in this game, and we define the PIC as the ratio between the maximal payment in a limit equilibrium of the game to the efficient cost.

Lemma 1 holds for matroids [9]. Thus, our proofs of Lemmas 2-4, which only rely on Lemma 1, hold for matroids as well. Similarly, Lemmas 5 and 6, Propositions 1 and 2, and Corollary 1 only rely on Lemmas 1-4. Thus, our characterization of limit equilibria of the Bertrand game holds for general matroid structures. In particular, Proposition 1 reveals that all limit equilibria (not only equilibria in which sellers never play weakly dominated strategies [10]) of the Bertrand game are efficient, independently of the number of edges for which a seller can bid. Proposition 2 reveals that, when each seller is the unique bidder for exactly one edge, then the buyer’s payment ranges from $\lambda(c)$ to $\mu(c)$. Thus, the PIC is $\frac{\mu(c)}{\lambda(c)}$ as well.

¹²It is well known that each network structure is a matroid.

¹³ $\Delta(e, \gamma, M)$, $E(c)$, and essential elements of E in a matroid structure can be defined as in Section 2.1.

A circuit of $M \equiv (E, F)$ is a subset of E not in F and such that all its proper subsets belong to F . A cost matrix $c \in \mathbb{R}_+^E$ satisfies the triangle inequality on M if for each circuit $T \subseteq E$, no edge $e \in T$ costs more than half the aggregate cost of T . One can easily see that Proposition 3 holds for matroids: given a matroid M and metric costs c , $PIC(M, c) \leq \text{dimension}(M)$. Our examples in Section 4.2 illustrate that the whole range of values for the PIC in the metric domain can be achieved with network structures.

6.3 The case of multiple ownership

A seller who is the exclusive bidder for an edge can be viewed as the owner of this edge. Our results depend crucially upon the assumption that no bidder owns more than one edge.

If a seller owns a set of mutually disconnected edges (no two edges share a node), our Theorem 1 still holds. Indeed it is straightforward to adapt the proof of Proposition 2 to show the following. If the buyer purchases γ and edge $e \in \gamma$ is sold by seller i , then the maximum price the buyer can pay for edge e in equilibrium is $\min\{c_f | f \in \Delta(e, \gamma, M), f \notin l(i)\}$. Because edges in $l(i)$ are mutually disconnected, $Ad(e, \gamma) \cap l(i) = \emptyset$, therefore the maximum equilibrium overhead for e is at most $\min\{c_f | f \in Ad(e, \gamma)\}$, and the two bounds for the maximal PIC, 2 and $2\frac{n-1}{n-2}$, still hold.

If sellers own sets of possibly connected edges, these bounds are not preserved.

Proposition 1 still holds: an optimal tree is purchased in any (limit) equilibrium. But if a seller owns several edges forming a *cut* (a set of edges intersecting every spanning tree), he has monopsony power so in our model he can extract infinite surplus.

Even if no seller owns a cut, the PIC can increase quite significantly. Here is an example.

Suppose $n = kp$ and m sellers can bid each for a different k -clique of edges. That is, seller 1 bids for all edges connecting nodes $1, \dots, k$, seller 2 for all edges connecting nodes $k+1, \dots, 2k$, and so on until seller m who bids for edges inside $(m-1)k+1, \dots, mk$. Other edges are covered by other sellers. Assume that the cost of any edge inside any clique is zero, and the cost of all other edges is 1. Then the profile of bids where all edges carry a price of 1 (so only the m clique-owners overcharge), and the line-tree $1-2-\dots-n$ is purchased, is a limit equilibrium where the PIC is $\frac{n-1}{m-1}$. If $k=2$ a clique is a single edge so we are in the situation of section 4, and indeed $\frac{n-1}{m-1} = 2\frac{n-1}{n-2}$. For larger k we have $\frac{n-1}{m-1} = \frac{km-1}{m-1} \geq k$, so the PIC increases linearly in the clique-size.

We conjecture that when the number n of nodes becomes large, and each seller is the exclusive bidder for at most a k -clique, then the maximum PIC converges to k .

Appendix

Proof of Lemma 2. Let $M \equiv (E, F)$ and $c \in \mathbb{R}_+^E$. Let $\gamma \in \Gamma_M \setminus \Gamma_M(c)$. We prove that there is $e \in \gamma$ and $f \in E \setminus \gamma$ such that $c_f < c_e$ and $\gamma - e + f \in \Gamma_M$. Let $\gamma' \in \Gamma_M(c)$. Since $\gamma' \notin \Gamma_M(c)$, then $\gamma'(c) < \gamma(c)$. By Statement 1 in Lemma 1, there is a bijection $f : \gamma \setminus \gamma' \rightarrow \gamma' \setminus \gamma$ such that for each $e \in \gamma \setminus \gamma'$, $\gamma - e + f(e) \in \Gamma_M$. Then, there is $e \in \gamma \setminus \gamma'$ such that $c_{f(e)} < c_e$, for otherwise $\gamma'(c) = \sum_{f \in \gamma \cap \gamma'} c_f + \sum_{f \in \gamma' \setminus \gamma} c_f \geq \sum_{f \in \gamma \cap \gamma'} c_f + \sum_{f \in \gamma \setminus \gamma'} c_f = \gamma(c)$. For such a $e \in \gamma \setminus \gamma'$, $c_{f(e)} < c_e$ and $\gamma - e + f(e) \in \Gamma_M$ as claimed. \square

Proof of Lemma 3. Let $M \equiv (E, F)$ and $c \in \mathbb{R}_+^E$. Let $\gamma \in \Gamma_M(c)$, $e \in \gamma$, and $c'_e < c_e$. Let $\gamma' \in \Gamma_M(c_{-e}, c'_e)$. Then $e \in \gamma'$, for otherwise $\gamma'(c_{-e}, c'_e) = \gamma'(c) < \lambda_E(c)$. We prove that $\gamma \in \Gamma_M(c_{-e}, c'_e)$. Let $\gamma' \in \Gamma_M(c_{-e}, c'_e)$. Since $\gamma \in \Gamma_M(c)$, then $\sum_{f \in \gamma \setminus \gamma'} c_f \leq \sum_{f \in \gamma' \setminus \gamma} c_f$. Since $e \in \gamma \cap \gamma'$, then $\gamma(c_{-e}, c'_e) = c'_e + \sum_{f \in \gamma \cap \gamma', f \neq e} c_f + \sum_{f \in \gamma \setminus \gamma'} c_f \leq c'_e + \sum_{f \in \gamma \cap \gamma', f \neq e} c_f + \sum_{f \in \gamma' \setminus \gamma} c_f = \gamma'(c_{-e}, c'_e)$. Thus, $\gamma \in \Gamma_M(c_{-e}, c'_e)$. \square

Proof of Lemma 4. Let $M \equiv (E, F)$ and $c \in \mathbb{R}_+^E$.

1. Let $e \in E(c)$ and $c' \geq c$. We prove that $e \in E(c'_{-e}, c_e)$. Let $\gamma \in \Gamma_M(c)$ be such that $e \in \gamma$ and $\gamma' \in \Gamma_M(c'_{-e}, c_e)$. Suppose that $e \notin \gamma'$. By Statement 2 in Lemma 1, there is $f \in \gamma'$ such that $\{\gamma - e + f, \gamma' - f + e\} \subseteq \Gamma_M$. Then, $c_e \leq c_f$, for otherwise $\sum_{g \in \gamma - e + f} c_g < \gamma(c)$. Since $c_f \leq c'_f$, then $c_e + \sum_{g \in \gamma' - f} c'_g \leq \gamma'(c'_{-e}, c_e)$. Thus, $\gamma' - f + e \in \Gamma_M(c'_{-e}, c_e)$.

2. Let $e \in E(c)$ and suppose that for each $\gamma \in \Gamma_M(c)$, $e \in \gamma$. Let $c' \geq c$. Let $\gamma' \in \Gamma_M(c'_{-e}, c_e)$. We prove that $e \in \gamma'$. Let $\gamma \in \Gamma_M(c)$. Then, $e \in \gamma$. Suppose by contradiction that $e \notin \gamma'$. By Statement 2 in Lemma 1, there is $f \in \gamma'$ such that $\{\gamma - e + f, \gamma' - f + e\} \subseteq \Gamma_M$. Then, $c'_f \leq c_e$, for otherwise $\gamma'(c'_{-e}, c_e) \leq c_e + \sum_{g \in \gamma' - f} c'_g < \gamma'(c'_{-e}, c_e)$. Since $c_f \leq c'_f$, then $\gamma - e + f \in \Gamma_M(c)$. Thus, $e \in \gamma - e + f$. This is a contradiction. \square

Proof of Proposition 1.

Step 1. Fix a Bertrand game $(M, S, (c^i)_{i \in S})$, an ε -equilibrium q , and an edge $e \in \gamma(q)$ that is built by seller i , so $q_e^i = q_e^*$. Then $q_e^i \geq c_e^* - 2\varepsilon$ (recall that $c_e^* \equiv \min_{\{j: e \in l(j)\}} c_e^j$).

Assume to the contrary $q_e^i < c_e^* - 2\varepsilon$ and consider the following alternative bidding strategy \tilde{p}^i :

- $\tilde{p}_e^i \equiv c_e^i$;
- $\tilde{p}_f^i \equiv \max\{c_f^i, q_f^i\}$, for every $f \notin \gamma(q)$, and for every $f \in \gamma(q)$ that i does not build at q ;
- $\tilde{p}_g^i \equiv q_g^i - \frac{\varepsilon}{n-1}$ for every edge $g \in \gamma(q) \setminus \{e\}$ that i builds at q .

Let $F = \{g \in \gamma(q) \setminus \{e\} \mid i \text{ builds } g \text{ at } q\}$. By Lemmas 3 and 4, irrespective of the tie-breaking rule, seller i still builds at (q_{-i}, \tilde{p}^i) every edge in F : indeed his bids decrease strictly for those edges, and increase at least weakly for other edges, so each edge in F is essential at (q_{-i}, \tilde{p}^i) ; moreover, i is the unique winner

in each of these edges at (q_{-i}, \tilde{p}^i) . His total profit from edges in F decreases by less than ε , while he saves more than 2ε on e , whether or not he still builds e at (q_{-i}, \tilde{p}^i) . Finally he does not lose anything on other edges he may be called to build. Thus his net profit increases by more than ε , contradiction.

Step 2. Fix the game $(M, S, (c^i)_{i \in S})$, an inefficient spanning tree $\gamma \notin \Gamma_M(c^)$, and a limit equilibrium p such that γ is built in a sequence of ε -equilibria q where ε goes to zero. We derive a contradiction.*

By Lemma 2 there exist $g \in \gamma$ and $f \in \Delta(g, \gamma, M)$ such that $c_f^* < c_g^*$ and $\gamma - g + f$ is a cheaper spanning tree. Set $\delta = c_g^* - c_f^*$ and choose for some $\varepsilon < \frac{\delta}{5}$ an ε -equilibrium q where γ is built.

Let j be the seller chosen to build g at q , and i be such that $c_f^i = c_f^*$. Distinguish two cases.

Case 1: $i \neq j$. We construct an alternative strategy \tilde{p}^i as follows:

- $\tilde{p}_f^i \equiv q_g^j - \varepsilon$;
- $\tilde{p}_e^i \equiv \max\{c_e^i, q_e^i\}$, for every $e \notin \gamma$, and for every $e \in \gamma$ that i does not build at q ;
- $\tilde{p}_e^i \equiv q_e^i - \frac{\varepsilon}{n-1}$ for every edge $e \in \gamma$ that i builds at εp .

At (q_{-i}, \tilde{p}^i) the tree $\gamma - g + f$ is optimal so f is efficient; it is even essential by Lemma 3 and because $\gamma - g + f$ is also optimal at (q_{-i}, \bar{p}^i) where $\bar{p}^i \equiv q_g^j - \frac{\varepsilon}{2}$. As in Step 1 above, seller i still builds at (q_{-i}, \tilde{p}^i) every edge he was building at q , because their price has decreased strictly so they are part of all optimal trees at (q_{-i}, \tilde{p}^i) . His total profit decreases by at most ε for those edges. Moreover he now builds f as well (because every other bid for f is at least q_g^j), with a profit of more than 2ε : indeed by Step 1 and the definition of i we have

$$q_g^j - \varepsilon \geq c_g^j - 3\varepsilon \geq c_g^* - 3\varepsilon = c_f^i + \delta - 3\varepsilon > c_f^i + 2\varepsilon.$$

Finally he does not lose money for edges other than f that he was not building and may be called to build. Contradiction.

Case 2: $i = j$. The alternative strategy of seller i is now:

- $\tilde{p}_f^i \equiv q_g^i - \varepsilon$; $\tilde{p}_g^i = q_g^i$;
- $\tilde{p}_e^i \equiv \max\{c_e^i, q_e^i\}$, for every $e \notin \gamma$, and for every $e \in \gamma$ that i does not build at q ;
- $\tilde{p}_e^i \equiv q_e^i - \frac{\varepsilon}{n-1}$ for every edge $e \in \gamma \setminus \{g\}$ that i builds at q .

Once again, by Lemmas 3 and 4, any optimal tree at (q_{-i}, \tilde{p}^i) contains f , and includes all other edges different from g that i builds at q . So i still builds the latter edges, and f as well. Distinguish two cases. If the tree purchased at (q_{-i}, \tilde{p}^i) contains g , and i still builds g (for the same price), the same accounting argument shows that i 's profit increases by more than ε , and yields a

contradiction. If i does not build g any more (whether or not g is built at all), then his new profit on f exceeds his old profit (or loss) on g by more than 4ε :

$$\{(q_g^i - \varepsilon) - c_f^i\} - \{q_g^i - c_g^i\} = c_g^i - c_f^i - \varepsilon = c_g^* - c_f^* - \varepsilon = \delta - \varepsilon > 4\varepsilon$$

and the proof is complete. \square

Proof of Lemma 5. Let $M \equiv (E, F)$, $c \in \mathbb{R}_+^E$, and $e \in E(c)$.

(i) Let $\gamma \in \Gamma_M(c)$ such that $e \in \gamma$. We prove that $\mu_e(c) = \min\{c_f : f \in \Delta(e, \gamma, M)\}$.

We prove first that $\mu_e(c) \leq \min\{c_f : f \in \Delta(e, \gamma, M)\}$. Let $b \in \mathbb{R}_+$ be such that $b > \min\{c_f : f \in \Delta(e, \gamma, M)\}$ and $f \in \Delta(e, \gamma, M)$ such that $c_f < b$. Thus, $\lambda(c_{-e}, b) \leq \lambda(c) - c_e + c_f$. We claim that $e \notin E(c_{-e}, b)$. Suppose by contradiction that there is $\gamma' \in \Gamma_M(c_{-e}, b)$ such that $e \in \gamma'$. By Lemma 3, $\gamma' \in \Gamma_M(c)$. Thus, $\lambda(c_{-e}, b) = \lambda(c) - c_e + b$, implying $b \leq c_f$. This is a contradiction.

Let $b \equiv \min\{c_f : f \in \Delta(e, \gamma, M)\}$. We prove that $\gamma \in \Gamma_M(c_{-e}, b)$. Suppose by contradiction that $\gamma \notin \Gamma_M(c_{-e}, b)$. Let $c' \equiv (c_{-e}, b)$. By Lemma 2 there exists $g \in \gamma$ and $f \notin \gamma$ such that $c'_f < c'_g$ and $\gamma - g + f \in \Gamma_M$. Since $c_{-e} = c'_{-e}$, then $g = e$, for otherwise $c(\gamma - g + f) < \lambda(c)$. Thus, $c_f < b$ and $f \in \Delta(e, \gamma, M)$. This is a contradiction.

(ii) Clearly if $e \in E(c)$, then $\min\{c_f : f \in \Delta(e, \gamma, M)\} \geq c_e$. The first statement is then an easy consequence of Lemma 5 (i). Now, let $e \in E(c)$ and $\gamma \in \Gamma_M(c)$ be such that $e \in \gamma$ and e is non essential, i.e., there is $\gamma' \in \Gamma_M(c)$ such that $e \notin \gamma'$. By Statement 2 in Lemma 1 there is $f \in \gamma'$ such that $\{\gamma - e + f, \gamma' - f + e\} \subseteq \Gamma_M$. Since $c(\gamma) = c(\gamma') = \lambda(c)$, then $c_e = c_f$, for otherwise $c(\gamma - e + f) < \lambda(c)$ or $c(\gamma' - f + e) < \lambda(c)$. Thus, $\min\{c_f : f \in \Delta(e, \gamma, M)\} \leq c_e$ and $\mu_e(c) = c_e$.

(iii) From $\{\gamma, \gamma'\} \subseteq \Gamma_M(c)$ we get $\sum_{e \in \gamma \setminus \gamma'} c_e = \sum_{e' \in \gamma' \setminus \gamma} c_{e'}$. By Lemma 5 (ii) we have $c_e = \mu_e(c)$ for each $e \in \gamma \setminus \gamma' \cup \gamma' \setminus \gamma$. Combining these equalities

$$\sum_{e \in \gamma} \mu_e(c) - \sum_{e''} \mu_{e''}(c) = \sum_{e \in \gamma \setminus \gamma'} \mu_e(c) - \sum_{e' \in \gamma' \setminus \gamma} \mu_{e'}(c) = \sum_{e \in \gamma \setminus \gamma'} c_e - \sum_{e' \in \gamma' \setminus \gamma} c_{e'} = 0,$$

as stated.

For the second statement, we fix an edge $e \in E$, $b \in \mathbb{R}_+$, and two costs c, c' such that $c \leq c'$. By Lemma 4, if $e \in E(c_{-e}, b)$, then $e \in E(c'_{-e}, b)$. By (1) this implies that $c \rightarrow \mu_e(c)$ is weakly increasing for e fixed. Thus $b \rightarrow \mu(c_{-e}, b)$ is weakly increasing on any interval where a given tree remains optimal. At the boundary between two such intervals, we just proved that $\mu(c_{-e}, b)$ is the same when computed for either tree. This implies the desired monotonicity. \square

Proof of Proposition 2. Let $(M, S, (c^i)_{i \in S})$ be a network procurement Bertrand game.

Statement i) Let p be a limit equilibrium of $(M, S, (c^i)_{i \in S})$. We prove that $\lambda(c^*) \leq \lambda(p^*) \leq \mu(c^*)$. Let $q(\varepsilon)$ be a sequence of ε -equilibria that as $\varepsilon \rightarrow 0$, $q(\varepsilon) \rightarrow p$. By Step 1 in the proof of Proposition 1, if edge e is selected at $q(\varepsilon)$, i.e., $e \in \gamma(q(\varepsilon))$, and is built by agent i , then $q(\varepsilon)_e^i \geq c_e^* - 2\varepsilon$. Therefore if e is selected at p (i.e., $e \in \gamma(q(\varepsilon))$) for a sequence ε going to zero and such that

$\gamma(q(\varepsilon))$ is constant) and built by i , we have $p_e^i \geq c_e^*$. This proves that the buyer pays at least $\lambda(c^*)$.

Next we pick an arbitrary $\varepsilon > 0$, an ε -equilibrium q , and $e \in \gamma(q)$. We prove by contradiction that $q_e^* \geq \mu_e(c^*) + 3\varepsilon$ is impossible. This will imply $q_e^* \leq \mu_e(c^*)$ in the limit and conclude the proof.

Let $f \in E$ and $i \in N$ be such that $\mu_e(c_e^*) = c_f^* = c_f^i$, where $f \in \Delta(e, \gamma(q), M)$. Because he does not build e at q . Consider the switch by i to the bid $\tilde{p}_f^i \equiv c_f^i + 2\varepsilon$, recalling that i bids only for edge f . The price of edge f is now strictly below that of e , so by Lemma 3 f is essential at (q_{-i}, \tilde{p}^i) . Seller i 's profit goes up from zero to 2ε , contradiction.

Statement ii). Fix c, γ , and p as in (2), and $\varepsilon > 0$. We check that the profile of bids $q_e \equiv \mu_e(c) - \varepsilon$ if $e \in \gamma$ and $q_e \equiv c_e$ otherwise, is an ε -equilibrium. Let $e \in \gamma$. By definition $e \in E(c_{-e}, \mu_e(c))$. By Lemma 3 applied recursively, $\gamma \subseteq E(c_{-\gamma}, (\mu_e(c))_{e \in \gamma})$. By Lemma 3 again, each edge in γ is essential at $(c_{-\gamma}, (\mu_e(c) - \varepsilon)_{e \in \gamma})$. Thus, $\Gamma_M(q) = \{\gamma\}$. Each seller $e \notin \gamma$ has zero profit, and cannot get positive profit by changing his bid. A seller $e \in \gamma$ has the net profit $\mu_e(c) - c_e - \varepsilon$, possibly negative but no less than $-\varepsilon$. If he raises his price by more than ε , he is no longer part of the optimal tree (or trees). We conclude that no deviation raises his profit by more than ε , as desired. \square

Proof of Lemma 6. Suppose $\lambda(c) = \mu(c)$, and for some stage of an instance of KA for c there is only one edge, say $e \in E$, that can be selected. Let γ be the tree constructed in such an instance of KA. Since for each $f \in \Delta(e, \gamma, M)$, $\gamma + f - e \in \Gamma_M$, then $c_e < \min\{c_f : f \in \Delta(e, \gamma, M)\}$. Thus, $\mu_e(c) > c_e$ and $\mu(c) > \lambda(c)$. Conversely, suppose that each stage of each instance of KA leaves two or more choices, and pick $e \in E(c)$. Then one can choose an instance of KA that always select an edge different from e , so there is $\gamma \in \Gamma_M(c)$ such that $e \notin \gamma$. \square

Proof of Proposition 3. Let $\gamma \in \Gamma_M(c)$ and $e \in \gamma$. Since no agent has monopoly power, there is $\gamma' \in \Gamma_M$ such that $e \notin \gamma'$. By Lemma 1, there is $f \in \gamma'$ such that $\{\gamma - e + f, \gamma' - f + e\} \subseteq \Gamma_M$. Since $\gamma + f \notin F$, then it contains a minimum cycle, $T \subsetneq \gamma$. Now, since $\gamma \in F$, then $f \in T$. Since $c \in T(M)$, then $c_f \leq \sum_{g \in \gamma} c_g = \lambda(c)$. Thus, $\mu(c) \leq (n-1)\lambda(c)$ and $PIC(M, c) \leq n-1$. \square

Proof of Theorem 1. The network structure is the complete graph $M \equiv (G, F)$. Fix $c \in T(M)$ and $\gamma \in \Gamma_M(c)$. For any $e \in \gamma$ and any pair of edges $\{f, g\} \subseteq G$ forming a triangle with e , the triangular inequality $c_g \leq c_e + c_f$ and Lemma 5 (i) imply $\mu_e(c) \leq c_e + \min\{c_f : f \in Ad(e, \gamma)\}$. Therefore $\mu(c) \leq \lambda(c) + \sum_{e \in \gamma} \min\{c_f : f \in Ad(e, \gamma)\}$, or equivalently,

$$\frac{\mu(c)}{\lambda(c)} \leq 1 + \frac{\sum_{e \in \gamma} \min\{c_f : f \in Ad(e, \gamma)\}}{\sum_{e \in \gamma} c_e}. \quad (5)$$

Thus it is enough to prove that for each $\gamma \in \Gamma_M$ and each $c \in \mathbb{R}_+^\gamma$,

$$\sum_{e \in \gamma} \min\{c_f : f \in Ad(e, \gamma)\} \leq \sum_{e \in \gamma} c_e \text{ if } n \text{ is odd,} \quad (6)$$

$$\sum_{e \in \gamma} \min\{c_f : f \in Ad(e, \gamma)\} \leq \frac{n}{n-2} \sum_{e \in \gamma} c_e \text{ if } n \text{ is even.} \quad (7)$$

We use the following notation. Let $E \subseteq G$ and $\sigma : \{1, \dots, |E|\} \rightarrow E$ be a bijection. The convex cone of costs on E ordered by σ is $K_\sigma(E) \equiv \{c \in \mathbb{R}_+^E : c_{\sigma(1)} \leq c_{\sigma(2)} \leq \dots \leq c_{\sigma(|E|)}\}$. Its canonical basis is the set $\{b^k \in \{0, 1\}^E : 1 \leq k \leq |E|, b_{\sigma(j)}^k = 1 \Leftrightarrow j \geq k\}$: each $c \in K_\sigma(\gamma)$ is a *non negative* linear combination of these vectors.

Fix c and a bijection σ such that $c \in K_\sigma(\gamma)$. The summations on the left and right of (6) and (7) commute with non negative linear combinations in $K_\sigma(\gamma)$. Indeed the minimum in $\theta(c) = \min\{c_f : f \in Ad(e, \gamma)\}$ is achieved at some $c_{\sigma(k)}$ where k is the same for all c in $K_\sigma(\gamma)$. Therefore $\theta(\alpha c + \alpha' c') = \alpha \theta(c) + \alpha' \theta(c')$ for $c, c' \in K_\sigma(\gamma)$, and $\alpha, \alpha' \geq 0$ as claimed.

Thus it is enough to prove (6) and (7) for the profiles b^k in the canonical basis of $K_\sigma(\gamma)$. As σ is arbitrary, this amounts to prove (6) and (7) for costs c such that $c_e \in \{0, 1\}$ for all $e \in \gamma$. It will be convenient to describe each such c by the subset $B^\gamma \subseteq \gamma$ of its “free” edges: $e \in B^\gamma \Leftrightarrow c_e = 0; e \in \gamma \setminus B^\gamma \Leftrightarrow c_e = 1$. With the notation $Ad(B^\gamma, \gamma) = \cup_{e \in B^\gamma} Ad(e, \gamma)$, we have then $\sum_{e \in \gamma} c_e = n - |B^\gamma| - 1$ and

$$\min\{c_f : f \in Ad(e, \gamma)\} = \begin{cases} 0 & \text{if } e \in Ad(B^\gamma, \gamma), \\ 1 & \text{otherwise.} \end{cases}$$

Thus $\sum_{e \in \gamma} \min\{c_f | f \in Ad(e, \gamma)\} = n - 1 - |Ad(B^\gamma, \gamma)|$. The proof of (6) and (7) amounts now to the following statement. For each $\gamma \in \Gamma_M$ and $B \subseteq \gamma$

$$|Ad(B, \gamma)| \geq |B| \text{ if } n \text{ is odd; } (n-2)|Ad(B, \gamma)| + 2(n-1) \geq n|B| \text{ if } n \text{ is even.} \quad (8)$$

Property (8) is clear if $B = \emptyset$ so we assume from now on that B is non empty. We say that B , a non empty subset of γ , is of “type 1” if $|Ad(B, \gamma)| \geq |B|$, and of “type 2” if $|Ad(B, \gamma)| = |B| - 1$.

We prove by induction on $n = |V|$ the following property \mathcal{P}_n , combining statements (9), (10), and (11):

$$|Ad(B, \gamma)| \geq |B| - 1 \text{ for any } B, \emptyset \neq B \subseteq \gamma \text{ (there is no “type 3”),} \quad (9)$$

$$B \text{ is of type 1 if all the edges incident to some node of } \gamma \text{ are outside } B, \quad (10)$$

$$B \text{ is of type 1 if it contains at least two adjacent edges of } \gamma. \quad (11)$$

The statement \mathcal{P}_3 is clear because $|Ad(B, \gamma)| = |B|$, whether B contains a single edge or $B = \gamma$. We assume $\mathcal{P}_3, \dots, \mathcal{P}_{n-1}$ and fix an arbitrary subset B of γ . We prove (10) first. Let i be a node of which all incident edges are outside B .

Case 1: Suppose first i is a leaf,¹⁴ so its unique incident edge e is also a leaf, and is not in B . Distinguish two subcases.

If $Ad(e, \gamma)$ contains at least one edge f in B , the inductive assumption (9) implies $|Ad(B, \gamma - e)| \geq |B| - 1$; on the other hand $Ad(B, \gamma) = Ad(B, \gamma - e) + e$,

¹⁴A node of degree 1 (a single edge in γ has this node as an end); we also call edge $e = ab$ a leaf of γ if either a or b is a leaf.

so B is of type 1. If $Ad(e, \gamma) \cap B = \emptyset$, pick one edge f in $Ad(e, \gamma)$ and consider the subtree $\gamma[f; e]$ away from e and with f as a leaf, i.e., generated by f and all the nodes i such that f is between i and e . By the inductive assumption (10), if B intersects $\gamma[f; e]$ we have $|Ad(B \cap \gamma[f; e], \gamma[f; e])| \geq |B \cap \gamma[f; e]|$. Summing up these inequalities over f in $Ad(e, \gamma)$ such that $B \cap \gamma[f; e] \neq \emptyset$, gives (10) because the sets $Ad(B \cap \gamma[f; e], \gamma[f; e])$ are pairwise disjoint and cover $Ad(B, \gamma)$ when f varies in $Ad(e, \gamma)$; similarly the sets $B \cap \gamma[f; e]$ cover B .

Case 2: Suppose i is of degree two or more. For each edge e incident to i , hence outside B , apply as above the inductive assumption (10) to the subtree $\gamma[e; i]$ away from i and with e as a leaf (if $B \cap \gamma[e; i] \neq \emptyset$), then sum up as above over all edges incident to i .

We prove (11) next. If B contains at least two adjacent edges of γ , pick a non trivial maximal subtree δ of γ , entirely in B and containing these two edges. If $\delta = \gamma$ property (11) is clear. Otherwise every edge f between δ and $B \setminus \delta$ is outside B , so we can again apply (10) to $\gamma[f; \delta]$ and sum up over f , because the sets $Ad(B \cap \gamma[f; \delta], \gamma[f; \delta])$ are pairwise disjoint. Thus $Ad(B, \gamma)$ contains at least $|B \cap \gamma \setminus \delta|$ edges outside δ . Moreover it also contains all edges in δ . This implies $|Ad(B, \gamma)| \geq |B|$ as desired.

Finally we prove (9). Suppose B is such that $|Ad(B, \gamma)| < |B|$. By (11) no two edges in B are adjacent, and by (10) every node is the endpoint of an edge in B . Thus every node is the end point of a unique edge in B , in other words the edges of B form a perfect matching of V . In particular n is even and $|B| = \frac{n}{2}$. Moreover $Ad(B, \gamma) = \gamma \setminus B$ hence $|Ad(B, \gamma)| = \frac{n}{2} - 1$, completing the proof of (9) and of \mathcal{P}_n .

This argument also shows that if B is of type 2 then n is even, and in this case $|B| = \frac{n}{2}$, so that the right hand side inequality in (8) is an equality. The latter inequality holds if B is of type 1 as well, so the proof of (8) and of Theorem 1 is complete. \square

Proof of Proposition 4. We only have to prove statement *ii*). If n is even and the edges of γ contain a perfect matching B , then $|B| = \frac{n}{2}$, $|Ad(B, \gamma)| = \frac{n}{2} - 1$, and the same argument in the proof of Theorem 1 shows that for the metric cost $c_e = 0$ if $e \in B$, $= 1$ if $e \in E \setminus B$, (7) holds as an equality, or equivalently $PIC(M, c) = 2 \frac{n-1}{n-2}$.

Conversely, suppose for some $c \in T(V)$ we have $PIC(M, c) = 2 \frac{n-1}{n-2}$. Then the restriction of c to γ is non zero, because $c_e = 0$ for all $e \in \gamma$ implies $c \equiv 0$ and $\lambda(c) = \mu(c) = 0$, so $PIC(M, c) = 1$. Moreover (7) holds as an equality. Say $c \in K_\sigma(\gamma)$ then c is a positive linear combination of the cost vectors b^k in the canonical basis of this cone. Inequality (7) holds for each b^k , therefore it must be an equality for at least one non zero b^k (one with a strictly positive coefficient in the decomposition of c). The cost vector b^k takes only the values 0, 1, so we can represent it as above by the set B of its free edges; now equality in (7) amounts to

$$(n-2)|Ad(B, \gamma)| + 2(n-1) = n|B|.$$

As $b^k \neq 0$, B is a strict subset of γ , $|B| < n-1$. This implies $|Ad(B, \gamma)| < |B|$,

i.e., B is of type 2. We can repeat the argument three paragraphs earlier showing that B is a perfect matching of V . \square

Proof of Proposition 5 Fix $M = (E, F)$ with $|E| = |G| - 1$, $c \in T(M)$, and $\gamma \in \Gamma_M(c)$. We prove that $\sum_{e \in \gamma} \mu_e(c) \leq 3 \sum_{e \in \gamma} c_e$. First, suppose that for each two adjacent edges in γ , say $\{e, f\}$, the edge that completes the triangle containing e and f , belongs to E . The same argument that applies when E contains all possible edges (Theorem 1) gives:

$$\sum_{e \in \gamma} \mu_e(c) \leq 2 \frac{n-1}{n-2} \sum_{e \in \gamma} c_e.$$

Our claim follows because $2 \frac{n-1}{n-2} \leq 3$ for $n > 3$.

Suppose now that the edge that completes the triangle in $\{f, g\} \equiv \gamma_1 \subseteq \gamma$ is not in E . Set $f = ab, g = bd$, and let A be the set of edges that connect γ_1 with $\gamma \setminus \gamma_1$, i.e.,

$$A \equiv \{e \in \gamma \setminus \gamma_1 : \text{at least one node of } e \text{ is } a, b, \text{ or } d\}.$$

We distinguish two cases.

Case 1: A contains a single edge h . Let $m \equiv \left(\sum_{e \in \gamma} c_e \right) - c_h - c_f - c_g$. Then $\gamma \setminus \{e, f, g\}$ is a tree spanning the $n - 3$ nodes in $V \setminus \{a, b, d\}$, and no edge is missing in the corresponding graph. As $n \geq 6$ we can apply Theorem 1 to deduce $\sum_{e \neq f, g, h} \mu_e(c) \leq 3m$. Thus

$$\begin{aligned} \sum_{e \in \gamma} \mu_e(c) &\leq \mu_f(c) + \mu_g(c) + \mu_h(c) + 3m \\ &\leq 3(c_f + c_g + c_h) + 3m = 3 \sum_{e \in \gamma} c_e. \end{aligned}$$

Case 2: $|A| \geq 2$. Suppose first that γ can be partitioned into two subtrees γ_f and γ_g such that $f \in \gamma_f, g \in \gamma_g, |\gamma_f| \geq 2$, and $|\gamma_g| \geq 2$. Then Theorem 1 implies

$$\sum_{e \in \gamma} \mu_e(c) \leq 3 \sum_{e \in \gamma_f} c_e + 3 \sum_{e \in \gamma_g} c_e = 3 \sum_{e \in \gamma} c_e.$$

Now if γ cannot be partitioned into two subtrees γ_f and γ_g as above, all edges in A must have the same end-node i in γ_1 (it could be a, b , or d). One can partition γ into subtrees rooted at i and such that one of these subtrees is γ_1 . Each such subtree (except for γ_1 if $i = b$) is rooted at i , and has a leaf edge whose outer node is i . Let $\gamma_1, \gamma_2, \dots, \gamma_k$ be this partition, with $k \geq 3$ because $|A| \geq 2$. There are two sub-cases.

Sub-case 1. n is odd (so $|\gamma|$ is even). Then $\gamma \setminus \gamma_1$ is a tree with an even number of edges. Let $m \equiv \left(\sum_{e \in \gamma} c_e \right) - c_f - c_g$. Let $x \equiv c_{e^*} = \min_{e \in A} c_e$, so $x \leq \frac{m}{2}$ because $|A| \geq 2$. On the other hand $\gamma \setminus \gamma_1$ has an even number of edges and each edge connecting two nodes in the node set of $\gamma \setminus \gamma_1$ is in E , therefore Theorem 1 gives $\sum_{e \neq f, g} \mu_e(c) \leq 2m$. Finally $\mu_f(c) \leq c_f + c_g + c_{e^*}$ by triangularity and the fact that there is an edge (in E) forming a cycle with

f, g, e^* (if $i = a, d$) or with f, e^* (if $i = b$). Collecting those facts we compute now

$$\begin{aligned} \sum_{e \in \gamma} \mu_e(c) &\leq \mu_f(c) + \mu_g(c) + 2m \leq 2(c_f + c_g + x) + 2m \\ &\leq 3m + 2(c_f + c_g) \leq 3 \sum_{e \in \gamma} c_e. \end{aligned}$$

Sub-case 2. n is even (so $|\gamma|$ is odd). Then, there must be at least one $l \in \{2, \dots, k\}$ such that $|\gamma_l|$ is odd. Say $|\gamma_2|$ is odd, then $\bigcup_{l \geq 3} \gamma_l$ is a non-empty ($k \geq 3$) tree with an even number of edges. Applying the case $|A| = 1$ to $\gamma_1 \cup \gamma_2$, and Theorem 1 to $\gamma_3 \cup \dots \cup \gamma_k$, we have

$$\begin{aligned} \sum_{e \in \gamma} \mu_e(c) &= \sum_{e \in \gamma_1 \cup \gamma_2} \mu_e(c) + \sum_{e \in \gamma_3 \cup \dots \cup \gamma_k} \mu_e(c) \\ &\leq 3 \sum_{e \in \gamma_1 \cup \gamma_2} c_e + 2 \sum_{e \in \gamma_3 \cup \dots \cup \gamma_k} c_e \leq 3 \sum_{e \in \gamma} c_e. \end{aligned}$$

□

Proof of Proposition 6. For any $c \in U(M)$, $\gamma \in \Gamma_M(c)$ if and only if

$$\text{for each } f = ab \notin \gamma, c_f = \max_{e \in [a, b]_\gamma} c_e, \quad (12)$$

(recall that for an arbitrary c , optimality of γ implies $c_f \geq \max_{e \in [a, b]_\gamma} c_e$).

Pick $e \in \gamma$ and $f \equiv ab \in \Delta(e, \gamma)$: there is at least one edge $g \in Ad(e, \gamma) \cap [a, b]_\gamma$, and (12) implies $c_f \geq \max\{c_e, c_g\}$; conversely, for each $g \in Ad(e, \gamma)$ the edge f forming a triangle with e and g is in $\Delta(e, \gamma)$ and $c_f = \max\{c_e, c_g\}$. By Lemma 5 (i),

$$\mu_e(c) = \max\{c_e, \min\{c_f : f \in Ad(e, \gamma)\}\}. \quad (13)$$

Thus, for each $c \in U(M)$ we have $\mu(c) = \lambda(c) + \sum_{e \in \gamma} \min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\}$. Thus,

$$PIC(M, c) = 1 + \frac{\sum_{e \in \gamma} \min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\}}{\lambda(c)}. \quad (14)$$

Statement ii): n is even. As $U(M) \subseteq T(M)$, the “only if” statement follows from statement *ii)* in Lemma 5. Conversely if γ contains a perfect matching B , the cost function $c_e = 0$ for all $e \in B$; $c_e = 1$ for all $e \in \gamma \setminus B$ is clearly ultrametric, and as we checked immediately after stating Proposition 6, $PIC(M, c) = 2 \frac{n-1}{n-2}$.

Statement i): For “if,” suppose node a is a leaf of γ , and the leaf edge $f = ab$ is adjacent (in γ) to a single edge $g = bd$. It is easy to check that the cost function: $c_e = 1$ if e connects $\{a, b\}$ to $V \setminus \{a, b\}$, $c_e = 0$ otherwise, is ultrametric. Clearly $\lambda(c) = 1$, and (13) implies $\mu_f(c) = \mu_g(c) = 1$. Thus, $PIC(M, c) = 2$.

For “only if”, assume that for some $c \in U(M)$ we have $PIC(M, c) = 2$. First we mimic the argument in the proof of Proposition 4 showing that such a c can be chosen with $c_e \in \{0, 1\}$ for all $e \in \gamma$. By (14) and Theorem 1 we have for all $c \in U(M)$

$$\sum_{e \in \gamma} \min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\} \leq \sum_{e \in \gamma} c_e, \quad (15)$$

with equality if and only if $PIC(M, c) = 2$. The latter implies that the restriction of c to γ , denoted c as well, is non zero, and if $c \in K_\sigma(\gamma)$, it is a positive

linear combination of cost vectors b^k also in $K_\sigma(\gamma)$ with all coordinates 0 or 1. Moreover inequality (15) is linear in this cone, so for at least one of the vectors b^k (15) is an equality. And b^k can be extended (uniquely) to an ultrametric cost on E .

Next, as in the proof of Theorem 1, we represent our 0,1- cost function c by the set B of its free edges. For any $e \in \gamma$, we have $\min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\} = 1$ if e is an isolated edge in B , i.e., it is not adjacent to any other edge in B ; and $\min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\} = 0$ for any other edge in γ . Let B^* be the subset of isolated edges in B : equality in (15) amounts to $|B^*| = |\gamma \setminus B|$. We prove by contradiction below that at least one $e \in B^*$ is adjacent to exactly one edge in $\gamma \setminus B$: this edge is the announced leaf edge of which the inner end point has degree two.

Clearly each $e \in B^*$ is adjacent to at least one edge in $\gamma \setminus B$. Suppose each $e \in B^*$ is adjacent to at least two edges in $\gamma \setminus B$. Notice that an edge in $\gamma \setminus B$ is adjacent to at most two edges in B^* . Consider the bipartite graph with nodes the elements of B^* on the left side, and those of $\gamma \setminus B$ on the right side (the same number of nodes on each side). It is a simple matter to check that if each left node is connected to at least two right nodes, and each right node to at most two left nodes, the bipartite graph contains a cycle, which in turn gives a cycle in γ , contradiction. \square

Proof of Proposition 7.

ii) By the proof of Lemma 5 (i), for any c and $e \in E(c)$ we have $\lambda(c_{-e}, \infty) - \lambda(c) = \mu_e(c) - c_e$. Thus if the buyer chooses $\gamma \in \Gamma_M(c)$, she pays $\lambda(c) + \sum_{e \in \gamma} (\mu_e(c) - c_e) = \mu(c)$. by Lemma 5 (iii), this payment does not depend on the choice of γ .

i) Fix a profile of bids $(c_{l(i)}^i)_{i \in S} \in \mathbb{R}_+^S$, a tree $\gamma \in \Gamma_M(c^*)$, and an edge \bar{e} . For agent i such that $l(i) = \bar{e}$, we have $\lambda(c^*(-i)) - \lambda(c^*) \leq \lambda(c_{-\bar{e}}^*, \infty) - \lambda(c^*) = \mu_{\bar{e}}(c^*) - c_{\bar{e}}^*$. And for any such i who does not win edge \bar{e} we have $\lambda(c^*(-i)) - \lambda(c^*) = 0$. Finally $\mu_{\bar{e}}(c^*) - c_{\bar{e}}^* = 0$ if $\bar{e} \notin \gamma$. Summing up over all edges shows that total payment (4) is bounded by $\mu(c^*)$. \square

References

- [1] N. Ailon, M. Charikar, Fitting tree metrics: Hierarchical Clustering and Phylogeny, *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 2005.
- [2] A. Archer, E. Tardos, Truthful mechanisms for one-parameter agents, *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, 2001.
- [3] A. Archer, E. Tardos, Frugal path mechanisms, *Proceedings of the 13th Annual ACM-SIAM Symposium on Discret Algorithms*, 2002, pp. 991-999.
- [4] S. Bikhchandani, S. de Vries, J. Schummer, R. Vohra, An ascending Vickrey auction for selling bases of a matroid, Mimeo, 2010.

- [5] G. Bergantiños, J. Vidal-Puga, A fair rule in minimum cost spanning tree problems, *J. Econ. Theory* 137, 2007, pp. 326-352.
- [6] A. Bogomolnaia, H. Moulin, Sharing a minimal cost spanning tree: beyond the folk solution, *Games Econ. Behav.* 69, 2010, pp. 238-248.
- [7] T. Borgers, Iterated elimination of dominated strategies in a Bertrand-Edgeworth model, *Review of Economic Studies* 59, 1992, pp. 163-176.
- [8] R. Brânzei, S. Moretti, H. Norde, S. Tijs, The P-value for cost sharing in minimum cost spanning tree situations, *Theory and Decision*, 56, 2004, pp. 47-61.
- [9] R. A. Brualdi, Comments on bases in dependence structures, *Mull. Austral. Math. Soc.*, 1969, pp. 161-167.
- [10] N. Chen, A. Karlin, Cheap labor can be expensive, *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2007, pp. 707-715.
- [11] E. Elkind, A. Sahai, K. Steiglitz, Frugality in path auctions, *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2004, pp. 701-709.
- [12] D. Fudenberg, D. Levine, Limit games and limit equilibria, *J. Econ. Theory* 38, 1986, pp. 261-279.
- [13] J. Green, J.J. Laffont, *Incentives in public decision making*, North-Holland, New York, 1979.
- [14] J. Hirschleifer, J. Riley, *The analytics of uncertainty and information*, Cambridge Surveys of Economic Literature, Cambridge University Press, 1992.
- [15] N. Immorlica, D. Karger, E. Nikolova, R. Sami, First-price path auctions, *Proceedings of the 6th ACM Symposium on Electronic Commerce*, 2005, pp. 203-212.
- [16] E. Koutsoupias, C. Papadimitriou, Worst case equilibria, *Proceedings of the 16th Symposium on Theoretical Aspects of Computer Science*, 1999, pp. 404-413.
- [17] A. Karlin, D. Kempe, T. Tamir, Beyond VCG: Frugality of truthful mechanisms, *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 2005.
- [18] J. Kruskal, On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Am. Math. Soc.* 7, 1956, pp. 48-50.
- [19] N. Nisan, T. Roughgarden, E. Tardos, V. Vazirani, (Eds.), *Algorithmic Game Theory*, Cambridge University Press, 2007.

- [20] H. Norde, S. Moretti, S. Tijs, Minimum cost spanning tree games and population monotonic allocation schemes, *European Journal of Operations Research* 154, 1, 2001, pp. 84-97.
- [21] L. Simon, W. Zame, Discontinuous games and endogenous sharing rules, *Econometrica* 58, 1990, pp. 861-872.
- [22] W. Sharkey, Network models in economics. In: Ball et al. (eds.) *Handbooks of Operations Research and Management Science*, pp. 713-765. New York: Elsevier,1995.
- [23] K. Talwar, The price of truth: frugality in truthful mechanisms, *Proceedings of the 21st Annual Symposium on Theoretical Aspects of Computer Science*, 2003, pp. 608-619.
- [24] J. Tirole, *The Theory of Industrial Organization*, The MIT Press, Cambridge, London, 1988.