PRICING WITH MARKUPS IN INDUSTRIES WITH INCREASING MARGINAL COSTS

JOSÉ R. CORREA*, NICOLÁS FIGUEROA**, AND NICOLÁS E. STIER-MOSES***

* *Departamento de Ingeniería Industrial, Universidad de Chile*
jcorrea@dii.uchile.cl

** *Centro de Economía Aplicada, Departamento de Ingeniería Industrial, Universidad de Chile*
nicolasf@dii.uchile.cl

*** *Graduate School of Business, Columbia University*
stier@gsb.columbia.edu

Abstract. We study a game that models a market in which heterogeneous producers of perfect substitutes make pricing decisions in a first stage, followed by consumers that select a producer that sells at lowest price. As opposed to Cournot or Bertrand competition, producers submit a price function to the market, which maps their production level to a price. Solutions of this type of models are normally referred to as supply function equilibria, and the most common application is in electricity markets. In our model, producers face increasing marginal production costs and, in addition, cost functions are proportional to each other, and their magnitude depend on the efficiency of each particular producer. In this context, we prove necessary and sufficient conditions for the existence of equilibria in which producers use supply functions that replicate their cost structure.

We then specialize the model to monomial cost functions with exponent equal to $q > 0$, which allows us to reinterpret the simple supply functions as a markup charged on top of the production cost. We prove that an equilibrium for the markups exists if and only if the number of producers in the market is strictly larger than $1 + q$, and if an equilibrium exists, it is unique. The main result for monomial cost functions is to establish that the equilibrium is nearly efficient when the market is competitive. Here, an efficient assignment is one that minimizes the total production cost, ignoring payments because they are transfers within the system. The result holds because when there is enough competition, markups are bounded, thus preventing prices to be significantly distorted from costs.

Finally, we focus on the case when unit costs are linear functions on the production quantities. This simplification allows us to refine the previous bound by establishing an almost tight bound on the worst-case inefficiency of an equilibrium. This bound is a subproduct of an algorithm that we design to find such equilibrium. The bound states that when there are two equally-efficient producers and possibly other less efficient ones, the production cost under an equilibrium is at most 50 percent worse than the optimal one, and the worst-case gap between the two assignments decreases rapidly as competition increases. For instance, for three similarly-efficient producers plus perhaps other less efficient ones, the inefficiency is below 6.2 percent.

Keywords. Imperfect Competition, Supply Function Equilibrium, Pricing, Game Theory, Allocation Efficiency.

JEL Classification. C61, C72, D43, L11, Q41.

Date: December, 2008.

1. Introduction

Traditional *Cournot* or *Bertrand* market-competition models consider that producers of perfect substitutes of a good decide the quantity they are going to produce or the price at which they are going to sell their production (Mas-Colell, Whinston, and Green 1995). Strategic consumers learn these decisions and decide whether to buy or not and from whom to buy. In practice, however, firms can use more flexible strategies, submitting *functions* which map the quantity produced to prices. One can interpret these functions as either supply functions or price functions, which are actually the inverse of each other. This seminal model was put forward and studied by Klemperer and Meyer (1989) and its outcomes have been referred to as *supply function equilibria*.

While Klemperer and Meyer emphasized the importance of these strategies in environments with uncertainty, supply function equilibria are relevant to various production and service industries even when uncertainty is not explicitly modeled. The most obvious example is the case of electricity markets. Here, firms have to submit an actual supply function to a regulating agency. The agency dynamically adjusts prices to the market-clearing ones, and producers have to sell the quantity they specified for the resulting price. A second example is given by the revenue management procedures widely utilized by the airline industry to postpone pricing decision as much as possible. Prices are revised daily according to the best forecast of demand available, which is of course dynamic because new information is collected as time goes by. This pricing strategy can be encoded in a supply function because the firm adjusts the sale price according to the quantity demanded. Finally, in the consulting industry, although the hourly rate of consultants in a project may be quoted before the project starts, the firm has some flexibility when deciding whether some tasks are part of the project or not. This decision may depend on the overall workload in the consulting firm.

We consider the case of an industry with an arbitrary number of asymmetric and strategic firms that produce perfect substitutes of a good. Moreover, we assume that producers have decreasing returns to scale and use a similar "technology" although some may be more efficient than others. A typical example of this is electricity generation, where firms use their efficient generators first, and turn on their less efficient ones only when the demand is high enough to deplete the capacity of the more efficient generators. Firms make pricing decisions forecasting the demand they will face under each combination of supply functions offered by the different producers. In a second phase, consumers learn the price functions chosen during the first phase and, by an unspecified learning process, converge to an equilibrium in which they select producers selling at lowest prices. We assume that consumers are small enough so they act as price takers, which simplifies the second-stage game. The demand is inelastic and publicly known. This approximation is particularly good in the case electricity markets since the game is played very often and therefore the short-term demand is not sensitive to prices and can be inferred by the producers. Although it seems to be possible to relax the assumptions on the consumer market (i.e., price-taking consumers and elastic demand), we leave those extensions for follow-up work.

An equilibrium is not necessarily efficient, meaning that it need not minimize the total production cost because of the presence of negative externalities. A natural question in this model is to study the inefficiency induced by the existence of producers with market power. Obviously, we expect producers to obtain positive, and potentially big, profits at equilibrium, but in the context of inelastic supply this does not necessarily mean inefficiency because price functions can be overstated with respect to cost functions. Indeed, distortion in prices may lead to too many or too few consumers choosing a particular producer, thus inflating the total cost for the economy. The question is of particular importance for a regulator, who is particularly interested to know how big this inefficiency can be for a big class of instances, since he does not have precise information about the cost structure of the firms. We consider then a worst-case type of analysis. For a big set of market structures, characterized only by the competitiveness of the industry, what is the worst possible ratio between an equilibrium allocation of production and an efficient one? This ratio has been referred to as the *Price of Anarchy*, its analysis was initiated by Koutsoupias and Papadimitriou (1999), and since then has been studied in several games relevant to computer science and operations research. The ratio quantifies the efficiency-loss at equilibrium relevant to an extremely risk averse regulator, who only knows the general structure of the market, and not the particular cost structure of each firm. Our paper focuses in the pricing question by adding a first stage to the game. The main conclusion is that the distortion in prices created by firms acting strategically can have an impact in the efficiency of an equilibrium, but this impact is limited.

We study the welfare implications of imperfectly-competitive markets where a great degree of flexibility is given to the producers. When the worst-case ratio between the cost of the equilibrium allocation and the efficient one is small, a planner can be sure that, independently of the details of the market structure, there is no big loss in welfare due to market power. Since this is true even if the planner does not even have a Bayesian estimation of the cost structure of the firms, it can be concluded that there is no benefit in going to great lengths to acquire that information. However, when the risk-averse planner lacks more detailed information and the worst-case ratio between the cost of the equilibrium allocation and the efficient one is large, this indicates that he should gather more details of the market structure. The additional information could be used to refine the analysis, to design a better mechanism of competition or to constrain the strategies that producers are allowed to use.

Let us refer to the per-unit cost of producer a by $u_a(x_a)$, where x_a is the production quantity of the firm. As mentioned, firms are heterogenous but face cost functions with a similar structure, thus we consider that cost functions have the form $u_a(x_a) := c_a u(x_a)$ and are parameterized with a single number c_a. In this paper we shall assume that producers face increasing marginal costs. Note that when industries have decreasing marginal costs, equilibria are fully efficient since all consumers purchase from the most efficient firm (there are no capacity constraints) and this coincides with the socially-optimal assignment. Hence, having an arbitrary number of firms, heterogeneity among firms but similar cost functions, and increasing marginal costs provides a setting that is more general than was previously studied while keeping the question of understanding the efficiency-loss at equilibrium relevant.

We start by analyzing the existence of equilibria in a game where producers are constrained to choose supply functions among a family parameterized by one parameter. The supply functions in this family, which we refer to as "simple," replicate the production costs of the industry. Indeed, firms choose β_a and bid the supply function $S_a(p) := \beta_a u^{-1}(p)$, potentially charging significantly above their marginal cost. We prove that an equilibrium exists if there are enough producers, where the threshold depends on the function u. Moreover, the equilibrium is monotonic in the sense that more efficient firms bid higher production quantities (for each given price) and capture a bigger market share. When the number of firms is too low, an equilibrium fails to exist because the amount of competition is not enough to curb sale prices and prevent firms from overcharging. In this situation, a best response to the prices of other firms is to always charge a little more, making the actual strategy space unbounded and thus preventing the existence of a fixed point.

To provide further support of the single-parameter bidding in the form of simple supply functions, we also prove that this equilibrium is immune to arbitrary deviations: firms cannot increase their profits by deviating and choosing an increasing supply function that is non-simple. Therefore, the equilibrium supported by simple supply functions is also an equilibrium for the larger strategy-space of increasing functions. From a practical point of view, an equilibrium where producers adopt price functions that imitate the shape of their production costs is also justified by the widely used practice of setting prices equal to the cost plus a fixed margin. In fact, if we consider unit cost functions that are proportional the monomial $u(x) = x^q$ for a fixed $q > 0$, the supply functions of the form mentioned earlier can be reinterpreted as a price function that includes a markup charged over the production cost. In this case, the outcome of the game can be seen as a *markup equilibrium*, where firms declare a supply function of the form $S_a(p) = \beta_a p^{1/q}$. The case of $q = 1$ is particularly interesting because, besides being more tractable, it is relevant to practice. For instance, Baldick, Grant, and Kahn (2004) provide a detailed answer to the question of why it is relevant to consider linear cost and price functions in electricity markets.

To study the efficiency loss due to imperfect competition, we concentrate in the case of monomial cost functions. A good reason for doing that is that multi-stage games like the one we analyze in this paper frequently become intractable when more general cost functions are used (see, e.g., Engel, Fischer, and Galetovic 2004; Acemoglu and Ozdaglar 2007; Wichiensin, Bell, and Yang 2007; Xiao, Yang, and Han 2007; Weintraub, Johari, and Van Roy 2008). For general $q > 0$, we prove that a markup equilibrium exists if and only if the number of competitors is strictly larger than $1+q$. Moreover, whenever an equilibrium exists, it is unique. Note that if the decreasing returns to scale are steeper, more firms are needed for an equilibrium to exist. As marginal costs rise faster, firms can be less aggressive in trying to obtain large market shares, so more of them are needed to ensure that there is a best response with bounded prices. The only parameter that matters for the existence of an equilibrium is the number of competitors but, surprisingly, not the relation between their cost functions. However, this relation is crucial to understand the relative efficiency of the equilibrium. Our results concerning the price of anarchy depend on the *competitiveness* of the market, which is measured by the $\ell_{1/q}$-norm of the vector whose components are c_1/c_a, where c_1 is the lowest cost of any producer.

One of the main insights provided by our results is that an equilibrium assignment under monomial costs is nearly efficient whenever the market is sufficiently competitive. More precisely, we prove that the price of anarchy tends to 1 as the competitiveness of the market tends to infinity. Furthermore, even when the competition is scarce, the price of anarchy is bounded by a small constant. We provide an upper bound to the efficiency-loss for any competition level higher than $(1+q)^q$. Note that it is likely that an industry is competitive in practice, whenever entry costs are small, since a non-competitive industry with high profits will induce entry. A basic idea behind these results is to show that although the most efficient producers are more profitable than less efficient ones (since the market structure supports larger markups for them), when there is enough competition, markups are bounded and cannot be infinitely large. This prevents prices from having a completely different structure from costs, which implies that the optimal and the equilibrium assignments are similar.

For the case of $q = 1$, our results imply that an equilibrium exists if and only if there are three or more producers. We establish a bound on the worst-case inefficiency of an equilibrium that is almost tight in general, and tight for infinitely many competitiveness values. Numerically, the bound states that the production cost at equilibrium in the linear case is at most 50 percent worse than the optimal one for values of competitiveness higher than 2 (i.e. the price of anarchy is 3/2), and the worst-case gap between the two assignments decreases rapidly as competition increases. For instance, the inefficiency is already below 6.2 percent when the competition level equals 3. On the other hand, using this bound, we also construct (asymptotically) worst-case instances.

It is important to note, however, that the ratio between profits that firms experience at equilibrium and those that would be achieved if producers were *non-strategic* can be much larger than the ratio between the corresponding social costs. Interestingly, if all producers charge high markups, but the relation between their prices reflect their relative costs well, the profit gap can be large, but the efficiency gap may remain small. Indeed, for the case $q = 1$, we prove that when the competitiveness of an instance tends to 2, the ratio of the social cost at equilibrium to that of the social optimum remains bounded by 3/2 although markups and profits may tend to infinity.

The latter result comes as a subproduct of an algorithm that we design to find an equilibrium in the linear case. One of the main ideas behind it is to observe that we can normalize any instance so the equilibrium equations become significantly simpler. Although this normalization does not lead to a closed-form solution for the equilibrium, it does provide an efficient procedure to compute one. With this simplification we can write the price of anarchy of all instances with linear costs explicitly as a nonconvex program. This can be reduced further to a nonconvex integer programming problem with only six variables that has a very small integrality gap.

In the literature, supply function equilibrium models have also been used to analyze welfare in the context of mergers. For example, Akgün (2004) models a merger as the appearance of a new firm with a reduced cost function, since the new firm can avoid more easily the decreasing returns to scale by allocating production efficiently among different plants. His model specification considers linear unit cost functions and elastic demand, and in such a context, he finds that equilibria always exist and that mergers decrease total welfare but increase the profits of merging firms. McAfee and Porter (2009) consider a more general framework where producers *and* consumers have market power and submit supply/demand functions. They develop a new measure of concentration, which can be related to equilibrium markups, profits and market shares. Acemoglu, Bimpikis, and Ozdaglar (2009) consider a context in which firms compete for supply quantities and price. They study a question similar to ours in spirit, providing bounds for the efficiency loss of the *best* equilibria. Finally, closest to our work is the paper by Johari and Tsitsiklis (2008), who also consider supply function equilibria from the perspective of studying the worst-case inefficiency at equilibrium. They present interesting results for a model in which producers, also facing an inelastic demand of one unit, can only choose a parameter w that leads to a supply function of the form $S(p) = 1 - w/p$.

2. The Model

We consider a market in which producers in $A = \{1, \ldots, n\}$ sell identical goods. The per-unit production cost for each producer $a \in A$ is a function $u_a : \mathbb{R}_+ \to \mathbb{R}_+$ that depends on the production quantity $x_a \in \mathbb{R}_+$. We assume that all producers make use of similar 'technology' but some are more efficient than others. This is modeled by letting the cost function be equal to $u_a(x_a) := c_a u(x_a)$ where the function $u(x_a)$ is an indication of the industry's unit cost for production level x_a, and the parameter c_a measures the efficiency of producer $a \in A$. Without loss of generality, we order producers such that $c_1 \leqslant \cdots \leqslant c_n$.

We assume that u is increasing, differentiable, and bijective (i.e., evaluates to zero in zero and grows to infinity). Furthermore, we assume that $xu(x)$ is convex; in other words, industries face increasing marginal production costs, which is the case, e.g., when labor or production capacity is scarce or when there is congestion. Note that as u is bijective, its inverse u^{-1} is well defined. Furthermore, we make the technical assumption that $pu^{-1}(p)$ is convex, which is equivalent to the following two conditions: $u(x)u''(x) < 2(u'(x))^2$ and u^2/u' is increasing. Some examples of unit cost functions that satisfy all assumptions are polynomials with nonnegative coefficients, the exponential function $e^x - 1$, and the logarithmic function $\ln(1+x)$. Putting all the elements together, the total cost of producing x_a units of the good is $\kappa_a(x_a) := x_a u_a(x_a) = c_a x_a u(x_a)$, which is convex.

The game we consider consists of two stages. The first stage is a pricing game among the producers and the second stage is an assignment game among the consumers, who decide from whom to buy. In the first stage of the game, producers select a *supply function* $S_a(p)$ which maps the quantity they are willing to produce to the corresponding unit price and inform consumers of their supply function. Equivalently, producers could consider a *price function* $p_a(x) = S_a^{-1}(x)$ because, being the inverse of the supply function, it provides the same information. If producer a receives a total order of x_a units of the good from the consumers, each unit will be sold at price $S_a^{-1}(x_a)$. Our only assumption is that producers are limited to choose supply functions that are increasing. The supply function is chosen to balance the tradeoff between high per-unit revenue and low demand, or vice-versa. The goal of the producer is to maximize its profit, which equals $x_a(S_a^{-1}(x_a) - c_a u(x_a)) \geqslant 0$.

In the second stage, consumers select their suppliers. We assume that there are infinitely many consumers that require an aggregated demand of one unit. The assumption of a unit demand is just for simplicity; the structure of cost functions makes the choice of total demand irrelevant.

Furthermore, we assume that each consumer is small compared to the market—implying that all of them act as price takers—and that the demand is inelastic, although both assumptions can be relaxed. Consumers satisfy their demands with producers that sell at minimal price. Throughout the paper we represent the aggregate consumption decisions by the vector $x \in \mathbb{R}^A$, which can be viewed as the market shares of the producers.

2.1. **Nash Equilibria.** A *supply function equilibrium* of the producers' game is a vector of supply functions $(S_a)_{a\in A}$ that satisfies the Nash equilibrium condition: no producer can increase the profit by switching to another supply function when the rest of the supply functions are fixed. An equilibrium in the consumers' game is an assignment x^{NE} such that all consumers are buying at minimal price. These two games are played sequentially, making it a Stackelberg game in which producers are the leaders and consumers are the followers.

The second stage is simply a market-clearing game, in which the quantity x_a that producer a sells equals $S_a(p^*)$, where p^* is the market-clearing price. Since the total demand equals one, the market-clearing price p^* is the unique solution to the equation

$$\sum_{a\in A} S_a(p^*) = 1.$$

Anticipating the market-clearing process, producers choose supply functions to maximize their profits, which can be written as a function of p^*:

$$x_a^{\text{NE}}(S_a^{-1}(x_a^{\text{NE}}) - c_a u(x_a^{\text{NE}})) = S_a(p^*)(p^* - c_a u(S_a(p^*))).$$

Interestingly, the previous equation implies that an equilibrium is completely determined by the choice of supply functions. Indeed, the vector of functions determines a unique market-clearing price p^*, which in turn determines unique market shares $x_a^{\text{NE}} = S_a(p^*)$ for the producers. We thus have the following definition.

Definition 1. A vector of supply functions $(\bar{S}_a)_{a\in A}$ is a Nash equilibrium for the producers' game if and only if

$$\bar{S}_a(\bar{p})(\bar{p} - c_a u(\bar{S}_a(\bar{p}))) \geqslant S_a(p)(p - c_a u(S_a(p))), \tag{1}$$

for all $a \in A$ and for all increasing and concave supply functions $S_a(\cdot)$. Here, the market-clearing price at equilibrium $\bar{p}$ satisfies $\sum_{a\in A} \bar{S}_a(\bar{p}) = 1$, while the price p under the alternative strategy $S_a(\cdot)$ is the unique solution to $S_a(p) + \sum_{i\neq a} \bar{S}_i(p) = 1$.

2.2. **Optimal Assignment.** To quantify the quality of an assignment, we let the total production cost $C(x) := \sum_{a\in A} \kappa_a(x_a) = \sum_{a\in A} c_a x_a u(x_a)$ be our social cost function. This function captures whether consumers are matched to the most efficient producers. Notice that payments are not considered in this function because they are internal transfers that do not have an impact on the welfare of the system. The socially-optimal assignment x^{OPT} is the unique minimizer of $C(x)$ given by

$$x^{\text{OPT}} := \arg\min \left\{ C(x) : \sum_{a\in A} x_a = 1,\ x_a \geqslant 0 \right\}.$$

It is worth observing that an optimal assignment is achieved if producers charge their marginal cost. Indeed, producer a charges its marginal cost when its supply function is the inverse of $\kappa'_a(x_a)$. Indeed, $S_a(p) = (\kappa'_a)^{-1}(p)$ leads to a market-clearing price p^* satisfying $\sum_{a\in A}(\kappa'_a)^{-1}(p^*) = 1$. A simple calculation shows that the optimal assignment $x_a^{\text{OPT}} = (\kappa'_a)^{-1}(p^*)$. However, the distortion of costs introduced by supply functions at equilibrium can lead to an assignment that does not necessarily minimize $C(x)$ because of the existence of negative externalities. One of our goals is to find conditions under which the distortions of costs and the increase of the total cost at equilibrium are not too large.

3. Supply Function Equilibria

We now turn into characterizing the equilibria of the game played among producers. We first focus on simple supply functions that replicate the cost structure of producers. Restricting the search to this type of functions allows producers to greatly simplify the problem of finding an optimal supply function since, as opposed to searching in a space of infinite dimensions, they have to consider a single degree of freedom. We prove that an equilibrium exists as long as there are enough producers in the market and an equilibrium does not exist if there are too few producers in the market. Next, we justify this choice of supply functions by proving that the equilibria we find are still at equilibrium when producers can pick their functions from the set of increasing functions. The main implication of this result is that these equilibria, supported by simple supply functions, are robust. On the negative side, we prove that equilibria that are not supported by simple supply functions can be extremely different to what one should expect and therefore are not likely to arise in practice.

3.1. **Equilibria with Simple Supply Functions.** In this section, we assume that producers restrict their consideration to simple supply functions of the form $S_a(p) = \beta_a u^{-1}(p)$ for a parameter $\beta_a \geqslant 0$ chosen by them. Notice that these functions charge prices that replicate the cost structure of the industry. Indeed, the corresponding price function (i.e., the inverse of the supply function) is $p_a(x_a) = u(x_a/\beta_a)$ so producers select an amplification factor for the demand, and charge the industry's cost evaluated at this amplified demand. We characterize best responses on the space of β_a, and establish sufficient conditions for an equilibrium to exist and for it to not exist.

Under the assumption that all producers bid simple supply functions, let us consider arbitrary but fixed parameters $\beta_i > 0$ for producers $i \neq a$. First, we compute the best response β_a of producer a as a function on the values of β_i for the other producers $i \neq a$. To determine the market-clearing price p, notice that to satisfy the demand, we need that $\sum_{i\in A} \beta_i u^{-1}(p) = 1$, from where $p = u(1/\sum_{i\in A} \beta_i)$. Replacing this price into the profit function shown in (1), the profit of producer a equals

$$P_a(\beta_a) := \frac{\beta_a}{\sum_{i\in A}\beta_i}\left(u\left(\frac{1}{\sum_{i\in A}\beta_i}\right) - c_a u\left(\frac{\beta_a}{\sum_{i\in A}\beta_i}\right)\right). \tag{2}$$

The next proposition optimizes the profit function over β_a, and establishes that although it is possible that $P_a(\beta_a)$ is not concave, there is a unique solution to the maximization problem.

Proposition 1. *There is a unique solution to the problem* $\max\{P_a(\beta_a) : \beta_a \geqslant 0\}$*, which is achieved where* P_a' *vanishes.*

Proof. Note that P_a is continuous and differentiable, and satisfies that $P_a(0) = 0$ and $P_a(\beta_a) < 0$ for $\beta_a \to \infty$. Therefore, $P_a(\beta_a)$ is maximized in $(0, \infty)$ at a point where the derivative vanishes. To simplify notation, we make $B_{-a} := \sum_{i\neq a} \beta_i$. Using the change of variable $p := u(1/(\beta_a + B_{-a}))$, the producer a can be viewed as choosing the market-clearing price that corresponds to a value of β_a. Rewriting the profit of producer a as a function of this price p, we get that

$$\begin{aligned} P_a(p) :=& (1 - B_{-a}u^{-1}(p))(p - c_a u(1 - B_{-a}u^{-1}(p))) \\ =& p - B_{-a}pu^{-1}(p) - (1 - B_{-a}u^{-1}(p))c_a u(1 - B_{-a}u^{-1}(p)). \end{aligned} \tag{3}$$

Taking the derivative of this function, a necessary condition for a price p to be a best response is that $p + u'(u^{-1}(p))(u^{-1}(p) - 1/B_{-a}) = c_a g'(1 - B_{-a}u^{-1}(p))$, where we used $g(x) := xu(x)$ to simplify notation. Notice that the right-hand side is positive and decreasing. We are going to prove that the left-hand side is increasing whenever it is positive, which will imply the existence of a unique best response. Indeed, using the monotone change of variable $y = u^{-1}(p)$, the left-hand side as a function of y is $h(y) := u(y) + u'(y)(y - 1/B_{-a})$. Taking the derivative,

$$h'(y) = 2u'(y) + u''(y)\left(y - \frac{1}{B_{-a}}\right) > \frac{2u'(y)}{u(y)}\left(u(y) + u'(y)\left(y - \frac{1}{B_{-a}}\right)\right)$$

where the inequality follows from $y - 1/B_{-a} \leqslant 0$ and $u(x)u''(x) < 2(u'(x))^2$. In conclusion, $h(y) \geqslant 0$ implies that $h'(y) \geqslant 0$, proving the claim. Transforming back to the original variables, we have that the best response for producer a equals $\beta_a = 1/u^{-1}(p^*) - B_{-a}$. $\square$

An equilibrium in the space of simple supply functions is a vector $(\beta_a)_{a\in A}$ in which each producer plays a best response to the others' actions. Interestingly, an equilibrium has a very natural property, namely, more efficient producers obtain a larger β and thus capture a larger market share. This is summarized in the following corollary.

Corollary 2. *If producers i and j are such that $c_i < c_j$, then $\beta_i > \beta_j$ and $x_i > x_j$.*

Proof. It is immediate from the proof of Proposition 1 that an equilibrium simultaneously satisfies

$$p + u'(u^{-1}(p))(u^{-1}(p) - 1/B_{-a}) = c_a g'(1 - B_{-a}u^{-1}(p)) \qquad \text{for all } a \in A.$$

Thus, since g' is increasing, if $c_i < c_j$ we must have that $B_{-i} < B_{-j}$, or equivalently $\beta_i > \beta_j$. $\square$

Moreover, with Proposition 1 at hand, we are ready to establish a sufficient and a necessary condition on the existence of equilibria. The two conditions compare the number of producers that are competing in the market to the ratio of the marginal production cost and the unit production cost. We first establish that an equilibrium exists when there are enough producers. Afterwards, we complement the positive result by proving that an equilibrium cannot exist when there are too few producers. Although these two conditions are not complementary, for the case of monomials considered in Section 4, the two results provide a complete characterization of the existence of equilibria in the space of simple supply functions.

Theorem 3. *Assume the number of producers n participating in the market is strictly larger than $\tilde{n} := \max_{x\geqslant 0}(xu'(x) + u(x))/u(x)$. Then, the producers' game has an equilibrium where producers bid in the space of simple supply functions.*

Proof. Let $\Gamma(\beta) = (\Gamma_1(\beta_{-1}), \dots, \Gamma_n(\beta_{-n}))$ be the best response mapping, where $\Gamma_a(\beta_{-a})$ maps the supply functions of others to the best response of producer a. To prove that an equilibrium exists, we show that Γ maps a compact into itself and use Brower's fixed point theorem. In particular, we prove that best responses are bounded.

Let $B_{-a} := \sum_{i\neq a}\beta_i$. As in the previous proposition, we consider the change of variable $p = u(1/(B_{-a}+\beta_a))$, and the profit function $P_a(p) = (1 - B_{-a}u^{-1}(p))(p - c_a u(1 - B_{-a}u^{-1}(p)))$. Taking the derivative with respect to p, the condition that p is a critical point requires that

$$u^{-1}(p) + p(u^{-1})'(p) - 1/B_{-a} = c_a(u^{-1})'(p)(u(1 - B_{-a}u^{-1}(p)) + (1 - B_{-a}u^{-1}(p))u'(1 - B_{-a}u^{-1}(p))).$$

Using the definition of $\tilde{n}$ and the relation $(u^{-1})'(p) = 1/u'(u^{-1}(p))$, we have that

$$c_a\tilde{n}u(1 - B_{-a}u^{-1}(p)) \geqslant p + u'(u^{-1}(p))u^{-1}(p) - u'(u^{-1}(p))/B_{-a}\,.$$

Going back to variable β_a, and after some algebra,

$$c_a\tilde{n}\frac{u(\beta_a/(B_{-a}+\beta_a))}{u(1/(B_{-a}+\beta_a))} \geqslant 1 - \frac{\beta_a/B_{-a}}{B_{-a}+\beta_a}\,\frac{u'(1/(B_{-a}+\beta_a))}{u(1/(B_{-a}+\beta_a))}\,,$$

which becomes

$$\frac{u(\beta_a/(B_{-a}+\beta_a))}{u(1/(B_{-a}+\beta_a))} \geqslant \frac{B_{-a} - (\tilde{n}-1)\beta_a}{B_{-a}c_a\tilde{n}}$$

when we evaluate the hypothesis on $1/(B_{-a}+\beta_a)$. Furthermore, since the profit is nonnegative at equilibrium, $u\left(1/(B_{-a}+\beta_a)\right) - c_a u\left(\beta_a/(B_{-a}+\beta_a)\right) \geqslant 0$. Putting the two bounds together we conclude that

$$\frac{B_{-a}-(\tilde{n}-1)\beta_a}{B_{-a}c_a\tilde{n}} \leqslant \frac{u(\beta_a/(B_{-a}+\beta_a))}{u(1/(B_{-a}+\beta_a))} \leqslant \frac{1}{c_a}. \tag{4}$$

We use (4) to prove that best responses are bounded. Let $\bar{c} = \max_{i\in A} c_i$ and $\underline{c} = \min_{i\in A} c_i$. Let us consider that for all $i \neq a$ the parameter β_i is bounded by

$$0 < \varepsilon \leqslant \beta_i \leqslant M := \frac{1}{u^{-1}(\underline{c}u(1/n))} < \infty\,,$$

where $\varepsilon > 0$ is will be determined later. We have to prove that the best response β_a is bounded by the same constants. Let us first see the upper bound. From the second inequality in (4), the assumption $\beta_i \leqslant M$ for all $i \neq a$, and assuming that $\beta_a > M$, we have that:

$$\frac{1}{\underline{c}} \geqslant \frac{1}{c_a} \geqslant \frac{u(\beta_a/(B_{-a}+\beta_a))}{u(1/(B_{-a}+\beta_a))} \geqslant \frac{u(\beta_a/((n-1)M+\beta_a))}{u(1/(B_{-a}+\beta_a))} \geqslant \frac{u(1/n)}{u(1/\beta_a)}\,.$$

This inequality says that $\beta_a \leqslant 1/u^{-1}(\underline{c}u(1/n)) = M$ which is a contradiction. To prove the lower bound, we first let a be such that β_a is the smallest and assume that $\beta_a < \varepsilon$, then using $B_a \geqslant (n-1)\varepsilon$ we can write

$$\frac{1}{\bar{c}\tilde{n}} - \frac{(\tilde{n}-1)\beta_a}{(n-1)\varepsilon\bar{c}\tilde{n}} \leqslant \frac{u(\beta_a/(B_{-a}+\beta_a))}{u(1/(B_{-a}+\beta_a))} \leqslant \frac{u(\varepsilon/(B_{-a}+\varepsilon))}{u(1/(B_{-a}+\varepsilon))}\,. \tag{5}$$

The reminder of the analysis is divided into two cases: when $B_{-a} > (n-1)\sqrt{\varepsilon}$ and when $B_{-a} \leqslant (n-1)\sqrt{\varepsilon}$. In the former case we have

$$\frac{u(\varepsilon/(B_{-a}+\varepsilon))}{u(1/(B_{-a}+\varepsilon))} \leqslant \frac{u(\varepsilon/((n-1)\sqrt{\varepsilon}+\varepsilon))}{u(1/((n-1)M+\varepsilon))} \leqslant \frac{u(\sqrt{\varepsilon}/(n-1))}{u(1/(nM))}\,,$$

while in the latter

$$\frac{u(\varepsilon/(B_{-a}+\varepsilon))}{u(1/(B_{-a}+\varepsilon))} \leqslant \frac{u(1/n)}{u(1/(n\sqrt{\varepsilon}))}\,.$$

In both cases, since $u(0) = 0$ and $\lim_{x\to\infty} u(x) = \infty$ for $\varepsilon > 0$ small enough we have that (and this is how ε is defined).

$$\frac{u(\varepsilon/(B_{-a}+\varepsilon))}{u(1/(B_{-a}+\varepsilon))} \leqslant \frac{n-\tilde{n}}{\bar{c}\tilde{n}(n-1)}\,.$$

Putting this inequality back into (5) we obtain

$$\frac{1}{\bar{c}\tilde{n}} - \frac{(\tilde{n}-1)\beta_a}{(n-1)\varepsilon\bar{c}\tilde{n}} \leqslant \frac{n-\tilde{n}}{\bar{c}\tilde{n}(n-1)},$$

which is equivalent to $\beta_a \geqslant \varepsilon$, a contradiction.

Thus, we have proved that Γ is a continuous function that maps a compact set into itself. Brower's fixed point theorem implies that it has a fixed point, which is a Nash equilibrium. □

The following proposition looks at the opposite case and proves that an equilibrium does not exists if there are too few producers competing in the market.

Proposition 4. *If $n \leqslant \min_{x>0}(xu'(x)+u(x))/u(x)$, then the producers' game does not have an equilibrium where producers bid simple supply functions.*

Proof. Differentiating (2), the best response condition $P'_a(\beta_a) = 0$ can be written as

$$\left(u\Big(\frac{1}{\sum_{i\in A}\beta_i}\Big) - c_a u\Big(\frac{\beta_a}{\sum_{i\in A}\beta_i}\Big)\right)\sum_{i\neq a}\beta_i = \frac{\beta_a}{\sum_{i\in A}\beta_i}\left(u'\Big(\frac{1}{\sum_{i\in A}\beta_i}\Big) + c_a u'\Big(\frac{\beta_a}{\sum_{i\in A}\beta_i}\Big)\sum_{i\neq a}\beta_i\right). \tag{6}$$

Using that $c_a > 0$, the previous equation implies that

$$u\left(\frac{1}{\sum_{i\in A}\beta_i}\right)\sum_{i\neq a}\beta_i > \frac{\beta_a}{\sum_{i\in A}\beta_i}u'\left(\frac{1}{\sum_{i\in A}\beta_i}\right),$$

which added together for all $a \in A$ leads to

$$(n-1)u\left(\frac{1}{\sum_{i\in A}\beta_i}\right)\sum_{a\in A}\beta_a > u'\left(\frac{1}{\sum_{i\in A}\beta_i}\right).$$

This contradicts the assumption that $xu'(x) + u(x) \geqslant nu(x)$ for $x = 1/\sum_{i\in A}\beta_i$, and completes the proof. □

3.2. **Robustness of Equilibria.** Having established conditions under which a simple equilibrium exists, we now show that this equilibrium is robust. In this section, we prove that it is always a best response for producer a to bid a simple supply function of the form indicated above, independently of the supply functions of producers $i \neq a$. Hence, at the equilibrium characterized in Section 3.1, it is in the best interest of producers to maintain their choices of simple supply functions, even when they are allowed to bid an arbitrary supply function.

Note that Klemperer and Meyer (1989) gave a similar result for the special case of linear cost functions. They proved that among the many equilibria there exists one in which producers bid linear supply functions. Our result generalizes that both cost and supply functions have the same shape to arbitrary cost functions and to multiple and heterogenous producers.

Theorem 5. *Assume that each producer $i \in A$ bids a simple supply function of the form $S_i(p) = \beta_i u^{-1}(p)$ for a $\beta_i > 0$ of their choice. Take an arbitrary producer $a \in A$. If these supply functions are at equilibrium in the space of simple supply functions, then S_a is also a best response function if producer a chooses supply functions from a strategy space consisting of all nondecreasing functions.*

Proof. Let us consider that supply functions are fixed for producers $i \neq a$ and focus on producer $a \in A$.When computing a best response, producer a solves

$$\max_{S_a(\cdot)}\left\{S_a(p)(p - c_a u(S_a(p))) : \sum_{a\in A} S_a(p) = 1\right\}. \tag{7}$$

Since supply functions of others are fixed, the problem of producer a is equivalent to choosing the market-clearing price $p \in [0, \bar{p}_{-a}]$, where $\bar{p}_{-a}$ is the market-clearing price when producer a does not participate, i.e., $\sum_{i\neq a} S_i(\bar{p}_{-a}) = 1$. Indeed, any p in that interval can be achieved, and given the supply functions of the other producers, the market share at the market-clearing price chosen by producer a is determined by $S_a(p) = 1 - \sum_{i\neq a} S_i(p)$. Therefore, (7) is equivalent to $\max\{P_a(p) : p \in [0, \bar{p}_{-a}]\}$, where $P_a(p)$ is defined as in (3). Proposition 1 implies in particular that there is a unique global maximizer p^* of $P_a(p)$, which is interior because $P_a(0) \leqslant 0$, $P_a(\bar{p}_{-a}) = 0$ and $P'_a(\bar{p}_{-a}) < 0$.

The space of simple supply functions is rich enough to achieve price p^* because the only condition needed is that the market share x_a at price p^* equals $1 - \sum_{i\neq a} S_i(p^*)$. Indeed, the original S_a optimizes β_a among all nonnegative values and hence is a best response to the others' supply functions. □

A similar argument can be applied to any equilibrium to show that the shape of the supply function outside the market-clearing price is irrelevant. The next result proves that an equilibrium with arbitrary supply functions, can be restated using simple supply functions plus some additive constants.

Corollary 6. *Assume that each producer $i \in A$ bids a supply function $S_i(p)$ that is nondecreasing and differentiable. If $(S_i)_{i\in A}$ is at equilibrium in the space of supply functions, then this equilibrium is outcome-equivalent to one where supply functions have the form $\tilde{S}_i(p) = \gamma_i + \beta_i u^{-1}(p)$ for γ_i and β_i chosen by each producer.*

Proof. Let p^* be the market-clearing price under the equilibrium we consider. Using the argument in the proof of Theorem 5, the fact that $S_a(p^*)$ is a solution to (7), the equilibrium condition can be restated as

$$p^* \in \arg\max \left\{ \Big(1 - \sum_{i\neq a} S_i(p)\Big)\Big(p - c_a u\big(1 - \sum_{i\neq a} S_i(p)\big)\Big) : p \in [0, \bar{p}_{-a}] \right\}. \tag{8}$$

Denoting the objective function by $H_a(p)$, we compute

$$H_a'(p) = 1 + c_a \sum_{i\neq a} S_i'(p)\left(\Big(1-\sum_{i\neq a} S_i(p)\Big)u'\big(1-\sum_{i\neq a} S_i(p)\big) + u\big(1-\sum_{i\neq a} S_i(p)\big)\right) - \sum_{i\neq a}\big(S_i(p) + pS_i'(p)\big). \tag{9}$$

Notice that $0 < p^* < \bar{p}_{-a}$ because $H_a(0) \leqslant 0$, $H_a(\bar{p}_{-a}) = 0$ and $H_a'(\bar{p}_{-a}) < 0$. Hence, the optimality of p^* implies that $H_a'(p^*) = 0$

Since the price p^* is optimal from the perspective of producer a, we must have that $H_a'(p^*) = 0$. This hints that a producer just needs to know $S_i(p^*)$ and $S_i'(p^*)$ for producers $i \neq a$ to know that p^* is the optimal choice of price; the values and derivatives of supply functions at other prices are irrelevant. Thus, producers need only two parameters to setup their supply functions and influence the decisions of others. Indeed, we can construct a new equilibrium based on supply functions of the form $\tilde{S}_a(p) = \gamma_a + \beta_a u^{-1}(p)$, with values of $\beta_a = S_a'(p^*)u'(u^{-1}(p^*))$ and $\gamma_a = S_a(p^*) - \beta_a u^{-1}(p^*)$. Notice that the parameters are chosen such that the new supply functions and their derivatives are equal to the original ones at the market-clearing prices. Replacing these new supply functions in (9), it is clear that p^* is still extremal in (8).

Proceeding in a similar way to Proposition 1, it is easy to observe that once all producers $i \neq a$ are bidding $\tilde{S}_i(p)$ then, $H_a'(p) = 0$ has a unique solution. The solution p^* maximizes the objective function $H_a(p)$ and, hence, is the unique global maximum in (8). In conclusion, the new supply functions support an equilibrium that is outcome-equivalent to the original one because the market-clearing price and market shares are the same in both. □

This result is significant since it implies that any equilibrium can be reinterpreted as an equilibrium in which players bid simple supply functions plus an additive constant. The consideration of additive constants, though, provides the basis for possibly very inefficient equilibria. Indeed, one can construct an example in which the first producer is much more efficient than the second one, and nevertheless, the second one gets almost all the market. This situation is clearly not desirable, and this provides another justification for having looked at equilibria where producers just consider bidding simple supply functions.

4. Monomial Cost Functions

In this section, we concentrate on monomial cost functions of the form $u(x) = x^q$, where $q > 0$ is a fixed real number. Hence, the total cost equals $\kappa_a(x_a) = c_a x_a^{1+q}$ which is a good first-order approximation to industries with increasing marginal costs. The simplification of cost functions allows us to further simplify the structure of supply functions because under monomial cost functions the producers' decisions can be viewed as selecting a fixed markup to be added to the production cost. Furthermore, we can characterize equilibria in a sharper way, which we use to provide more structure, to prove the uniqueness of equilibria, to bound the supply functions as a function on the competitiveness among producers, and to construct an efficient algorithm to compute equilibria in the case of linear cost functions.

First note that if u is a monomial, it satisfies all the assumptions required by the model (see Section 2). Applying Theorem 5 to monomial cost functions implies that the simple supply functions introduced in Section 3.1 are $S_a(p) = \beta_a p^{1/q}$. Therefore, the corresponding price function can be expressed as

$$p_a(x_a) = \frac{u(1/\beta_a)}{c_a} \cdot c_a u(x_a) = \alpha_a \cdot c_a x_a^q,$$

where we have separated the factor that encodes the producer's decision and denoted it by α_a. Notice that this factor α_a takes the form of a markup added to the production cost that is independent of the production quantity. From now on, we will consider the markups $(\alpha_a)_{a\in A}$ to be the strategic variables, and the vector of markups corresponding to a supply function equilibrium will be referred to as a *markup equilibrium*.

Reinterpreting Corollary 2 in the setting of markups, we know that at equilibrium the most efficient producers charge higher markups and that their market shares are larger. Intuitively, since the efficient producers know that consumers are going to buy regardless of the price, they can increase the price to a level similar to the less efficient ones. Hence, we have that $\alpha_1 \geqslant \cdots \geqslant \alpha_n$ and that $x_1^{\text{NE}} \geqslant \cdots \geqslant x_n^{\text{NE}}$.

4.1. **Optimal Assignment and Best Responses.** With the structure we put in place, we can obtain explicit formulas for the optimal assignment and for the unique assignment corresponding to a given vector of markups. Indeed, $C(x)$ is a convex function and all producers are active under both assignments. Let us start by providing a definition that quantifies the competitiveness of the market.

Definition 2. The competitiveness of an instance is $\sigma := c_1(\sum_{i\in A}(1/c_i)^{1/q})^q \geqslant 1$. In this case we say that the instance is σ-competitive. Note that σ is the $\ell_{1/q}$-norm of the vector whose components are c_1/c_i.

This definition quantifies the variability of the efficiency of the producers, relying on the structure of the instance and not on the equilibrium itself. For values of σ similar to one, there is a large gap between the most efficient and most inefficient producers, implying a rather monopolistic environment. Large values of σ translate into competitive economic environments in which there are either a large number of producers or all producers have similarly efficient production facilities.

With this definition, the first-order optimality conditions of the optimal assignment problem give that

$$x_a^{\text{OPT}} = \frac{(1/c_a)^{1/q}}{\sum_{i\in A}(1/c_i)^{1/q}} = \left(\frac{c_1}{c_a\sigma}\right)^{1/q} . \tag{10}$$

and the optimal social cost is

$$C(x^{\text{OPT}}) = \sum_{a\in A} x_a^{\text{OPT}} c_a u(x_a^{\text{OPT}}) = \frac{c_1}{\sigma} .$$

It also follows immediately that the equilibrium allocation, given a vector of markups $\vec{\alpha}$, is

$$x_a^{\text{NE}} = \frac{1/(\alpha_a c_a)^{1/q}}{\sum_{i\in A} 1/(\alpha_i c_i)^{1/q}} , \tag{11}$$

and under it, the total production cost equals

$$C(x^{\text{NE}}) = \sum_{a\in A} x_a^{\text{NE}} c_a u(x_a^{\text{NE}}) = \left(\frac{1}{\sum_{i\in A} 1/(\alpha_i c_i)^{1/q}}\right)^{1+q} \left(\sum_{a\in A} \frac{1}{(c_a)^{1/q}(\alpha_a)^{1+1/q}}\right) .$$

Equation (11) and the ordering of the market shares $x_1^{\text{NE}} \geqslant \cdots \geqslant x_n^{\text{NE}}$ imply that $\alpha_1 c_1 \leqslant \cdots \leqslant \alpha_n c_n$.

To obtain an optimal markup α_a, a producer needs to balance the tradeoff between charging high to increase revenue and charging low to increase sales. Producers achieve this by anticipating consumers decisions when maximizing their profit. Producer a finds α_a by solving $\arg\max_{\alpha_a \geqslant 1} P_a(\alpha_a)$, where the objective function captures its profit

$$P_a(\alpha_a) := (\alpha_a - 1)\, x_a^{\text{NE}} c_a u(x_a^{\text{NE}}) = c_a(\alpha_a - 1) \cdot \left(1 + \sum_{i \neq a} \left(\frac{\alpha_a c_a}{\alpha_i c_i} \right)^{1/q} \right)^{-(1+q)}$$

as a function of its own decision α_a, while others' markups α_i for $i \neq a$ are fixed. By Proposition 1, the optimal markup is characterized by the first-order conditions of the optimization problem. Then, the best response $\Gamma_a(\alpha_{-a})$ for producer $a \in A$ to a given vector of markups α_{-a} is given by the unique solution to the equation:

$$\alpha_a = 1 + q + q \frac{\alpha_a}{(c_a \alpha_a)^{1/q} \sum_{i \neq a} 1/(\alpha_i c_i)^{1/q}} . \tag{12}$$

It is interesting to interpret this best response function: it says that the optimal markup is to charge the marginal cost (the term equal to $1 + q$) plus another term that depends on the competition in the market.

For linear cost functions (i.e., when $q = 1$), α_a cancels out from the right-hand side so the expression for the optimal markups is explicit and in closed-form. In addition, the socially-optimal assignment can be written as $x_a^{\text{OPT}} = c_1/(c_a \sigma)$, where the competitiveness measure is $\sigma = c_1 \sum_{a \in A} 1/c_a \geqslant 1$. With this, the optimal social cost can be simplified to $C(x^{\text{OPT}}) = 1/\big(\sum_{a \in A} 1/c_a\big) = c_1/\sigma$. The assignment of consumers to producers at equilibrium satisfies that $x_a^{\text{NE}} = (\alpha_a c_a \sum_i 1/(\alpha_i c_i))^{-1}$. Finally, the markup equilibrium for the game can be characterized by the following system of nonlinear equations:

$$c_a \alpha_a = 2c_a + \frac{1}{\sum_{i \neq a} 1/(\alpha_i c_i)} \quad \text{for all } a \in A, \tag{13}$$

where the marginal price, the marginal cost and the markup, which depends only on the marginal prices of others, are written explicitly.

Unfortunately, we do not know how to solve the previous system and therefore we cannot compute the equilibrium directly, even for the case of linear unit costs. Then, to prove that a markup equilibrium $\vec{\alpha}$ exists for general monomials, we look for a fixed point of the mapping $\Gamma : \vec{\alpha} \to (\Gamma_a(\alpha_{-a}))_{a \in A}$.

4.2. **Characterization of Equilibria.** We now show that a unique equilibrium exists if and only if the number of producers exceeds $1 + q$, and then study some properties of the markups at equilibrium. We remark that Turnbull (1983) already showed that there is a unique supply function equilibrium in the space of linear supply functions, for the case of two producers with linear cost functions. The following result generalizes the uniqueness of simple equilibria to the case of monomial cost functions of an arbitrary degree $q > 0$, and to multiple and heterogenous producers.

Proposition 7. *If $u(x) = x^q$ for $q > 0$, then there exist equilibria if and only if $n > 1 + q$. Furthermore, if the equilibrium exists, it is unique.*

Proof. When $u(x) = x^q$, we have that $(xu'(x) + u(x))/u(x) = 1 + q$, making the conditions of Theorems 3 and 4 exact opposites. Hence, we have a necessary and sufficient condition for the existence of an equilibrium. We only need to show that whenever an equilibrium exist, it is unique.

Let $\vec{\alpha}$ be a markup equilibrium. Observe that if we replace all costs by μc_a, for a scaling factor $\mu \in \mathbb{R}_+$, $\vec{\alpha}$ still solves (12) for all $a \in A$. Letting the factor μ be $\big(\sum_{a \in A} 1/(\alpha_a c_a)^{1/q}\big)^q$, we can express (12) simply by

$$\alpha_a\left(1-\frac{q}{(c_a\mu\alpha_a)^{1/q}-1}\right)=1+q \quad \text{for all } a\in A. \tag{14}$$

Consider the left-hand side of the previous equation as a function of α_a, keeping μ fixed. When it is positive, the left-hand side is continuous and strictly increasing on α_a because both of its terms are strictly increasing. Thus, for a fixed μ there can be a single value of α_a that satisfies the equation.

If there were two different equilibria $\vec{\alpha}$ and $\vec{\alpha}'$, their corresponding scaling factors μ and μ' have to be different. But this is a contradiction since $c_a\mu\alpha_a$ in the solution of (14) is monotone on the value of μ. Indeed, increasing μ makes α_a decrease because of (14). In addition, letting $w=(c_a\mu\alpha_a)^{1/q}$, (14) also gives that $w=q\alpha_a/(\alpha_a-1-q)+1$. Since $\partial w/\partial\alpha_a=-q(1+q)/(\alpha_a-1-q)^2\leqslant 0$, we have that w is monotone on μ. Hence, $\sum_{a\in A}1/(\alpha_a\mu c_a)^{1/q}\neq\sum_{a\in A}1/(\alpha'_a\mu' c_a)^{1/q}$, which contradicts the definition of μ or μ'. □

The proof says that if we could guess the appropriate scaling factor μ for the costs, we could compute an equilibrium by solving (14) for all $a\in A$. Although this equation cannot be solved in closed form, we will use this idea to provide an algorithm for the linear case.

Observe that $\sigma>(1+q)^q$ implies that $n>1+q$, from where $\sigma>(1+q)^q$ is a sufficient condition for an equilibrium to exist. On the contrary, if there are $1+q$ producers with equal cost, we have that $\sigma=(1+q)^q$ and an equilibrium does not exist. Since we want to use the competitiveness σ of an instance to provide bounds on the markups and the inefficiency of the resulting equilibria, we are going to adopt the previous condition on σ, which is the tightest possible, to guarantee existence.

Proposition 8. *Assume that $q>0$ and $\sigma>(1+q)^q$. A markup equilibrium $\vec{\alpha}$ satisfies that $1+q\leqslant\alpha_a\leqslant(1+q)(\sigma^{1/q}-1)/(\sigma^{1/q}-1-q)$ for all $a\in A$.*

Proof. The lower bound is immediate after (12) because a producer will never charge less than the marginal cost. Let us bound the right-hand side of equation (12) for the producer applying the largest markup, which is the first as we observed before. By the previous corollary for all $a\in A$.

$$\alpha_a\leqslant\alpha_1=1+q+q\frac{\alpha_1}{(\alpha_1c_1)^{1/q}\sum_{i>1}1/(\alpha_ic_i)^{1/q}}\leqslant 1+q+q\frac{\alpha_1}{c_1^{1/q}\sum_{i\in A}1/c_i^{1/q}-1}=1+q+q\frac{\alpha_1}{\sigma^{1/q}-1}.$$

The upper bound follows after we factor α_1 out of the previous inequality. □

For example, for linear cost functions, we have that $2\leqslant\alpha_a\leqslant 2(\sigma-1)/(\sigma-2)$. In particular, if $\sigma=4$, we know that α has to be between 2 and 3. This formula can also be used to find the minimum competitiveness that must be present to guarantee that markups will not exceed a given number.

4.3. **Computation of Equilibria with Linear Cost Functions.** This section provides an algorithm that computes the unique markup equilibrium in the case of linear cost functions (i.e., when $q=1$). We assume that $n\geqslant 3$ because otherwise we know that an equilibrium cannot exist. As before, we will choose the scaling factor μ that will simplify calculations. Assume that $n\geqslant 3$ and that $\vec{\alpha}$ is a markup equilibrium. Let $\mu:=\sum 1/(\alpha_ac_a)>0$. Considering the instance with costs equal to μc_a and replacing $\alpha_a\mu c_a$ by a variable w in (12), we get that w satisfies $w^2-2(\mu c_a+1)w+2\mu c_a=0$. Solving this quadratic equation we get that $w=1+\mu c_a+\sqrt{1+(\mu c_a)^2}$, or equivalently

$$\alpha_a=1+1/(\mu c_a)+\sqrt{1+1/(\mu c_a)^2}. \tag{15}$$

Now that we know w, the condition that defines μ is

$$\sum_{a\in A}1/(1+\mu c_a+\sqrt{1+(\mu c_a)^2})=1. \tag{16}$$

We have just shown that existence of an equilibrium implies that the markup equilibrium has to satisfy (15), where μ is defined by (16). Notice that this characterization can also be used to prove existence of equilibrium, because it can be seen that (16) has a solution if and only if $n \geqslant 3$. More importantly, it can be used to provide an algorithm to compute the markup equilibrium.

Proposition 9. *An approximate markup equilibrium can be computed efficiently.*

Proof. Let μ be such that (16) holds. Clearly, we can find an approximation $\tilde{\mu}$ using binary search on (16), and then use it to compute an approximate markup equilibrium $(\tilde{\alpha}_a)_{a\in A}$ using (15) and its allocation

$$\tilde{x}_a^{\text{NE}} := \frac{1}{\tilde{\alpha}_a \tilde{\mu} c_a} = \frac{1}{1 + \tilde{\mu} c_a + \sqrt{1 + (\tilde{\mu} c_a)^2}}\,. \tag{17}$$

If $\tilde{\mu}$ approximates (16) within an additive $\varepsilon > 0$, then it is easy to see that $|\tilde{\mu} - \mu| \leqslant O(1)\varepsilon$ and also that $|\tilde{x}_a^{\text{NE}} - x_a^{\text{NE}}| \leqslant O(1)\varepsilon$. □

5. Analysis of Efficiency for Monomial Cost Functions

In this section, we analyze the efficiency-loss at the equilibrium $(\vec{\alpha}, x^{\text{NE}})$ of a game, when compared to the optimal assignment x^{OPT}. In other words, we provide bounds on the price of anarchy for the game, which quantifies the loss generated by the lack of coordination in the system (Koutsoupias and Papadimitriou 1999).

As we will see below, if we do not restrict the instances we consider, the assignment at equilibrium can be arbitrarily bad compared to the social optimum and the markups applied to costs can be arbitrarily large. This high inefficiency comes from instances where producers are extremely different. Suppose there is a very efficient producer and a very inefficient one. Although in an optimal assignment for the market most consumers buy from the efficient producer, at equilibrium the efficient producer will add a big markup to its cost to match the inefficient producer. Hence, as opposed to the optimal situation, the market shares at equilibrium will be comparable, making the social cost of an equilibrium much higher than that of an optimal assignment of consumers to producers.

For this reason, we will parametrize all the bounds on the efficiency-loss at equilibrium with respect to the competitiveness of the market, measured by σ. Recall that, by Proposition 7, $\sigma > (1+q)^q$ guarantees the existence and uniqueness of equilibrium, while an equilibrium may not exist when $\sigma = (1+q)^q$. Thus, we concentrate in the former case.

First, we develop a general upper bound for the case when production costs are described by monomial functions. For $\sigma \to \infty$ the upper bound approaches to 1, which proves that an equilibrium is almost optimal in a highly-competitive market. We then refine the upper bound in the case when unit costs are linear functions on the production quantity. The latter bound is almost tight, and allows us to show that the price of anarchy is exactly 3/2 for arbitrary values of $\sigma \geqslant 2$. Parametrizing it with fixed values of σ, it decreases rapidly when σ increases.

5.1. An Initial Upper Bound on the Inefficiency.

In this section we analyze instances with monomial cost functions of arbitrary degree $q > 0$ and provide bounds that are parametric on the competitiveness in the instance σ. Using the fact that markups at equilibrium cannot be arbitrarily large, we provide an initial upper bound on the worst-case production cost at equilibrium that is simple but loose.

Proposition 10. *Consider a markup equilibrium for a σ-competitive instance with monomial cost functions of degree $q > 0$. If $\sigma > (1+q)^q$, then the price of anarchy is bounded by $(\sigma^{1/q} - 1)/(\sigma^{1/q} - 1 - q)$.*

Proof. The assumption implies that an equilibrium x^{NE} exists. For any producer $a \in A$, (11) implies that $\sum_{i\in A} c_i\alpha_i(x_i^{\text{NE}})^{1+q} = c_a\alpha_a(x_a^{\text{NE}})^q$. Then,

$$(1+q)C(x^{\text{NE}}) \leqslant \sum_{i\in A} c_i\alpha_i(x_i^{\text{NE}})^{1+q} = c_a\alpha_a(x_a^{\text{NE}})^q = \Big(\sum_{i\in A}\frac{1}{(\alpha_i c_i)^{1/q}}\Big)^{-q} \leqslant \frac{(1+q)(\sigma^{1/q}-1)}{\sigma^{1/q}-1-q}C(x^{\text{OPT}})\,.$$

The two inequalities follow, respectively, from the lower and upper bounds on the markups given in Proposition 8. Therefore, $C(x^{\text{NE}}) \leqslant (\sigma^{1/q}-1)/(\sigma^{1/q}-1-q)\cdot C(x^{\text{OPT}})$ □

As an example, this bound evaluates to $(\sigma-1)/(\sigma-2)$ when production costs are linear and, in particular, to 3/2 when $\sigma = 4$. In words, although there is no coordination and producers and consumers maximize their individual utilities, the inefficiency generated by competition cannot be extremely large. The increase in social cost at equilibrium cannot be more than 50 percent of the optimal social cost.

The bound given by Proposition 10 follows from the bounds we have on the markups α_a. The following proposition significantly improves this result by working directly with the market shares x^{NE} at equilibrium. Although we do not know how to express x^{NE} in closed-form, we prove an upper and a lower bound on it and relax the equilibrium condition by considering a nonlinear programming problem that captures the essence of the calculation of the price of anarchy. The bound follows from setting the market shares to the worst-case value among those that are feasible.

Theorem 11. *Consider a markup equilibrium for a σ-competitive instance with monomial cost functions of degree $q > 0$. Assume that ℓ and u are two positive numbers such that $\ell(c_1/c_a)^{1/q} \leqslant x_a^{\text{NE}} \leqslant u(c_1/c_a)^{1/q}$. If $\sigma > (1+q)^q$, then the price of anarchy is bounded by*

$$(\ell\sigma^{1/q})^{1+q} + \sigma(1-\ell\sigma^{1/q})\,\frac{u^{1+q}-\ell^{1+q}}{u-\ell}\,.$$

Proof. Since we do not know how to characterize a markup equilibrium exactly, we will relax the requirement that market shares are at equilibrium and consider an arbitrary market share vector that satisfies the box constraints $\ell(c_1/c_a)^{1/q} \leqslant x_a^{\text{NE}} \leqslant u(c_1/c_a)^{1/q}$. To find an upper bound on the worst-case inefficiency of an equilibrium $C(x^{\text{NE}})/C(x^{\text{OPT}})$, we solve the following nonlinear programming problem:

$$\begin{aligned} \max\quad & \frac{\sigma}{c_1}\sum_{a\in A} c_a x_a^{1+q} \\ \text{s.t.}\quad & \sum_{a\in A} x_a = 1 \\ & \ell(c_1/c_a)^{1/q} \leqslant x_a^{\text{NE}} \leqslant u(c_1/c_a)^{1/q} \qquad a\in A. \end{aligned}$$

Considering slack variables z_a from the lower bound, any feasible solution can be written as $x_a = \ell(c_1/c_a)^{1/q} + z_a$, and the first constraint is equivalent to $\sum_a z_a = 1 - \ell\sigma^{1/q}$. Since $c_1 \leqslant \dots \leqslant c_n$, an optimal solution satisfies that

$$z_a = \begin{cases} 0 & \text{for } 1 \leqslant a < k \\ 1-\ell\sigma^{1/q} - \sum_{i\neq a} z_i & \text{for } a = k \\ (u-\ell)(c_1/c_a)^{1/q} & \text{for } k < a \leqslant n. \end{cases}$$

Here, k is determined so that x_k satisfies the box constraints. Evaluating the optimal objective value and using Newton's generalized binomial theorem (Graham, Knuth, and Patashnik 1994, p. 162), we get the a bound on the price of anarchy. In the following derivation, we also use the upper bound for z_a and the expression for their sum.

$$\begin{aligned}
\frac{\sigma}{c_1}\sum_{a\in A} c_a x_a^{1+q} =&\sigma \sum_{a\in A}\sum_{k=0}^{\infty} \tbinom{1+q}{k}\left(\tfrac{c_1}{c_a}\right)^{(1-k)/q} \ell^{1+q-k} z_a^k\\
=&\sigma \ell^{1+q}\sum_{a\in A}\left(\tfrac{c_1}{c_a}\right)^{1/q} + \sigma\sum_{a\in A} z_a \sum_{k=1}^{\infty}\tbinom{1+q}{k}\left(\tfrac{c_1}{c_a}\right)^{(1-k)/q}\ell^{1+q-k} z_a^{k-1}\\
\leqslant&(\ell\sigma^{1/q})^{1+q} + \sigma\sum_{a\in A} z_a\sum_{k=1}^{\infty}\tbinom{1+q}{k}\ell^{1+q-k}(u-\ell)^{k-1}\\
=&(\ell\sigma^{1/q})^{1+q} + \sigma(1-\ell\sigma^{1/q})\sum_{k=1}^{\infty}\tbinom{1+q}{k}\ell^{1+q-k}(u-\ell)^{k-1}\\
=&(\ell\sigma^{1/q})^{1+q} + \frac{\sigma(1-\ell\sigma^{1/q})}{u-\ell}\sum_{k=1}^{\infty}\tbinom{1+q}{k}\ell^{1+q-k}(u-\ell)^{k}\\
=&(\ell\sigma^{1/q})^{1+q} + \frac{\sigma(1-\ell\sigma^{1/q})}{u-\ell}\left(\left(\sum_{k=0}^{\infty}\tbinom{1+q}{k}\ell^{1+q-k}(u-\ell)^{k}\right)-\ell^{1+q}\right)\\
=&(\ell\sigma^{1/q})^{1+q} + \sigma(1-\ell\sigma^{1/q})\frac{u^{1+q}-\ell^{1+q}}{u-\ell}\,.
\end{aligned}$$

Since $\ell \leqslant u$ and $\ell\sigma^{1/q} \leqslant 1$, the bound is well defined. ☐

Interestingly, the upper bound converges to 1 for $\sigma \to \infty$, for any fixed value of q. This says that when competition is high, then equilibria are almost efficient. Unfortunately, for small σ (i.e., $\sigma \approx (1+q)^q$) the bound becomes rather loose; actually, it approaches infinity for $\sigma \to (1+q)^q$.

As an example, let us see that the framework put forward by the previous theorem can be used to provide a meaningful bound on the price of anarchy. Consider the case of a monomial cost function $u(x) = x^q$ for $q \geqslant 1$. To get lower and upper bounds on x_a^{NE}, we use Proposition 8. Applying (11) to the denominator of (12), we have that

$$x_a^{\text{NE}} = \frac{1}{1+q+\frac{q(1+q)}{\alpha_a-1-q}} = \frac{1}{1+q+\frac{(1+q)(\alpha_a c_a)^{1/q}}{\alpha_a}\sum_{i\neq a}\frac{1}{(c_i\alpha_i)^{1/q}}}.$$

Since $q \geqslant 1$, the previous expression is nondecreasing as a function of α_i for all $i \in A$. Thus, we can provide a lower bound using that $\alpha_a \geqslant 1+q$:

$$x_a^{\text{NE}} \geqslant \frac{1}{q+(\sigma c_a/c_1)^{1/q}} \geqslant \frac{1}{q+\sigma^{1/q}}\left(\frac{c_1}{c_a}\right)^{1/q}.$$

For the upper bound, (11) together with the bounds for α_a imply that

$$x_a^{\text{NE}} \leqslant \left(1+\frac{q}{\sigma^{1/q}-1-q}\right)^{1/q}\frac{1}{\sum_{i\in A}(c_a/c_i)^{1/q}} = \left(\frac{\sigma^{1/q}-1}{\sigma(\sigma^{1/q}-1-q)}\right)^{1/q}\left(\frac{c_1}{c_a}\right)^{1/q}.$$

Putting all together, we can set $\ell = 1/(q+\sigma^{1/q})$ and $u = ((\sigma^{1/q-1}-1/\sigma)/(\sigma^{1/q}-1-q))^{1/q}$ in the previous theorem to get a bound on the price of anarchy. In particular, when production costs are linear the bound is $(\sigma^2-\sigma-1)/(\sigma^2-\sigma-2)$, which evaluates to $11/10$ when $\sigma = 4$.

Notice that Theorem 11 generalizes Proposition 10 since using $\ell = 0$ and the previous u in the theorem yields the weaker proposition. Finally, note also that better upper and lower bounds on the market shares at equilibrium could be given if one iterates the best responses additional times. Instead of continuing in that direction, we shall focus on linear cost functions, and use the characterization of equilibria proposed previously to provide a bound that is almost tight in general and exactly tight for infinitely many values of σ. To provide a benchmark to evaluate the previously-cited value of 11/10, for $\sigma = 4$ the price of anarchy is approximately 1.027. Moreover, the price of anarchy is exactly 3/2 if one considers all instances with $\sigma \geqslant 2$.

5.2. **Tight Bounds using a Nonlinear Programming Formulation.** We now compute a tight bound on the price of anarchy for linear cost functions. Since a worst-case example obviously matches the best upper bound, our goal is to come up with the worst-case example among instances that are σ-competitive. In the context of computing the worst-case inefficiency of equilibria, this approach was pioneered by Johari and Tsitsiklis (2004). We characterize this value exactly for a fixed value of $\sigma \geqslant 2$ by optimizing over the values of c_a:

$$POA(\sigma) := \sup_{n \geqslant 3} \left\{ \max \ \left(\sum_{a=1}^{n} \frac{1}{c_a} \right) \left(\sum_{a=1}^{n} \frac{c_a}{(1 + c_a + \sqrt{1 + c_a^2})^2} \right) \right\} \tag{18a}$$

$$\text{s.t.} \ \sum_{a=1}^{n} \frac{1}{1 + c_a + \sqrt{1 + c_a^2}} = 1 \tag{18b}$$

$$c_1 \sum_{a=1}^{n} \frac{1}{c_a} = \sigma \tag{18c}$$

$$0 < c_1 \leqslant c_a \qquad \forall a \in \{2, \ldots, n\}. \tag{18d}$$

Here, (18b) guarantees that we can use the characterization given previously to compute equilibria, (18c) imposes that the instance is σ-competitive, and the objective function, which equals $C(x^{\text{NE}})/C(x^{\text{OPT}})$, computes the inefficiency of the markup equilibrium of the instance represented by the feasible solution. Notice that this allows us to compute the inefficiency of an equilibrium without explicitly computing the equilibrium. For a fixed $n \geqslant 3$, the maximum in (18) is attained. We will see that the supremum is attained at the limit when n grows to infinity. On the other hand, the constraint $c_1 > 0$ can be relaxed to $c_1 \geqslant 0$ because in a worst-case solution $c_1 \neq 0$; otherwise, $C(x^{\text{OPT}}) = C(x^{\text{NE}}) = 0$.

It is not clear how to solve the previous problem directly because it has a nonlinear objective and nonlinear constraints (and both are non-convex). To get around that, we will transform its variables and reformulate the problem. Indeed, consider variables $0 \leqslant y_a \leqslant 1$ for all $a \in A$ such that

$$y_a := 1 - \frac{2}{1 + c_a + \sqrt{1 + c_a^2}}\,. \tag{19}$$

Note that the inverse transformation is $c_a = 2y_a/(1 - y_a^2)$. Replacing the new variables in (18), we reformulate the problem as

$$\begin{aligned} POA(\sigma) = \sup_{n \geqslant 3} \ \max \ & \frac{\sigma}{4}\Big(\frac{1}{y_1} - y_1\Big)\Big(n + 2 - 2\sum_{a=1}^{n} \frac{1}{1 + y_a}\Big) \\ \text{s.t.} \ & \sum_{a=1}^{n} y_a = n - 2 \\ & \sum_{a=1}^{n} \Big(\frac{1}{y_a} - y_a\Big) = \sigma\Big(\frac{1}{y_1} - y_1\Big) \\ & 0 \leqslant y_1 \leqslant y_a \leqslant 1 \qquad \forall a \in \{2, \ldots, n\}. \end{aligned} \tag{20}$$

Here, the objective and constraints of the new formulation correspond to those of the original one. Thus, the formulations (18) and (20) coincide, and their solutions are in one-to-one correspondence.

Lemma 12. *The maximum in the subproblem of* (20) *is increasing with* n*.*

Proof. Increasing n to $n' > n$ increases the objective because given a solution with n components, one can obtain the same objective by setting the new $n' - n$ variables to 1. Indeed, all constraints are satisfied and the objective value does not change. Hence, the optimal value with n' components has to be larger. □

The argument in the proof also implies that variables y_a taking a value of 1 are not useful. Indeed, if for a given value of n, it is optimal to set some variables to 1, we can remove those producers without effecting the feasibility of the solution, nor its objective value.

Corollary 13. *There is a solution to* (20) *in which all variables y_a are strictly smaller than 1.*

Let us consider the following problem in which we take $n \geqslant 3$ and $0 < y_1 < 1$ fixed, and optimize over $y_2, \dots, y_n$. We characterize the structure of the optimal solution to this problem and optimize over n and y_1 afterwards to get the solution to (20).

$$\begin{aligned}
\min\ & \sum_{a=2}^{n} \frac{1}{1+y_a} \\
\text{s.t.}\ & \sum_{a=2}^{n} y_a = n - 2 - y_1 \\
& \sum_{a=2}^{n} \frac{1}{y_a} = \sigma\left(\frac{1}{y_1} - y_1\right) + n - 2 - \frac{1}{y_1} \\
& y_1 \leqslant y_a \leqslant 1 \qquad \forall a = 2, \dots, n.
\end{aligned} \tag{21}$$

We denote the dual variables of this subproblem by λ, μ, ℓ_a and u_a, in the order in which they appear in the formulation above. We use the standard KKT conditions to characterize the optimal solution to this problem (see, e.g., Nocedal and Wright 2006, p. 321). Indeed, if a vector $y = (y_2, \dots, y_n)$ is optimal, then when the gradients of the active constraints at y are linearly independent, y verifies the KKT conditions

$$\frac{-1}{(1+y_a)^2} + \lambda - \frac{\mu}{y_a^2} + u_a - \ell_a = 0 \qquad \forall a = 2, \dots, n. \tag{22}$$

We use the previous condition to prove that an optimal solution can have at most two different values.

Proposition 14. *The variables at an optimal solution to* (21) *can take at most two values in the open interval* $(y_1, 1)$.

Proof. Assume vector $y = (y_2, \dots, y_n) \in [y_1, 1)^{n-1}$ is an optimal solution to (21). We can assume that there are at least three different values larger than y_1, because otherwise we have the claim. We refer to three of those values with y_i, y_j and y_k, ordered from low to high. The KKT conditions hold for y because the gradients of the active constraints at y are linearly independent. Indeed, the gradients of the two equality constraints are $(1, \dots, 1)$ and $(-1/y_2^2, \dots, -1/y_n^2)$, and those of the variables with values equal to y_1 are $(0, \dots, 0, -1, 0, \dots, 0)$ with the -1 in the position corresponding to the variable. The constraints of variables with value different from y_1 are not active so they are not considered. Considering the variables y_i and y_j only, the restricted gradients are $(1, 1)$, $(-1/y_i^2, -1/y_j^2)$, and $(0, 0)$. Because the first two are linearly independent, all the vectors are linearly independent and, thus, y_i, y_j and y_k satisfy (22). Using the complementary slackness property for these variables and solving for λ, we have that

$$\lambda = \frac{1}{(1+y_i)^2} + \frac{\mu}{y_i^2} = \frac{1}{(1+y_j)^2} + \frac{\mu}{y_j^2} = \frac{1}{(1+y_k)^2} + \frac{\mu}{y_k^2}.$$

We focus first on y_i and y_k. Solving for μ and plugging the result back in, we get that

$$\lambda = \frac{\frac{1}{(1+y_i)^2} - \frac{1}{(1+y_k)^2}}{1 - \left(\frac{y_k}{y_i}\right)^2} + \frac{1}{(1+y_k)^2} = \frac{1}{y_i^2 - y_k^2}\left[\left(\frac{y_i}{1+y_i}\right)^2 - \left(\frac{y_k}{1+y_k}\right)^2\right].$$

Taking the derivative with respect to y_k, we see that the right-hand side of the previous equation is increasing on y_k. After some algebra, the derivative is positive if and only if

$$\left(\frac{y_k}{1+y_k}\right)^2 - \left(\frac{y_i}{1+y_i}\right)^2 > \frac{y_k^2 - y_i^2}{(1+y_k)^3},$$

which holds because $y_i < y_k$. Now, doing the same calculation with y_i and y_j provides us with another λ. But that cannot happen because y_j and y_k satisfy the KKT conditions. ☐

The previous result implies that an optimal solution to (21) has the structure $(y_1, \ldots, y_1, y_i, \ldots, y_i, y_j, \ldots, y_j)$, where y_1 is repeated k_1 times and y_i is repeated k_i times and y_j is repeated $k_j = n - 1 - k_1 - k_i$ times. Hence, we can reformulate the previous problem as:

$$\begin{aligned}
\min\ & \frac{k_1}{1+y_1} + \frac{k_i}{1+y_i} + \frac{k_j}{1+y_j} \\
\text{s.t. }\ & k_1 y_1 + k_i y_i + k_j y_j = n - 2 - y_1 \\
& \frac{k_1}{y_1} + \frac{k_i}{y_i} + \frac{k_j}{y_j} = \sigma\left(\frac{1}{y_1} - y_1\right) + n - 2 - \frac{1}{y_1} \\
& k_1 + k_i + k_j = n - 1 \\
& y_1 \leqslant y_i \leqslant y_j \leqslant 1, k_1, k_i, k_j \in \mathbb{N},
\end{aligned} \tag{23}$$

where y_i, y_j, k_1, k_i, and k_j are the variables, and y_1 and n are fixed. We can thus conclude the following result.

Theorem 15. *The price of anarchy for a given value of $\sigma > 1$ is given by:*

$$\begin{aligned}
POA(\sigma) = \sup\ & \frac{\sigma}{4}\left(\frac{1 - y_1^2}{y_1}\right)\left(n + 2 - \frac{2(k_1+1)}{1+y_1} - \frac{2k_i}{1+y_i} - \frac{2k_j}{1+y_j}\right) \\
s.t.\ & 0 \leqslant y_1 \leqslant 1, n \geqslant 3, n \in \mathbb{N} \\
& (y_i, y_j, k_1, k_i, k_j) \textit{ solves problem } (23) \textit{ for } y_1 \textit{ and } n.
\end{aligned} \tag{24}$$

Observe that, although the previous problem only has six variables, it is a nonconvex integer programming problem. Hence it may be difficult to solve in closed form. Nevertheless, we can compute fairly tight lower and upper bounds to this problem, which we can express in closed-form as a function of $\sigma \geqslant 2$. The upper bound follows from a relaxation of problem (23), while the lower bound comes from feasible solutions for the same problem.

The following proposition shows that relaxing the integrality of the k variables results in at most two values of y in an optimal solution, as opposed to Proposition 14 that proves that there are at most three values of y.

Proposition 16. *Consider a problem similar to (23) in which k_1, k_i and k_j are relaxed to the nonnegative reals. There is an optimal solution satisfying that $y_i = y_j$.*

Proof. Consider a relaxation of a the subproblem of (23) in which k_1 is fixed, and k_i and k_j are variables, and the three are reals numbers. If k_i or k_j is zero, we can assume the claim because the corresponding value of y is not used. To get a contradiction, assume that an optimal solution satisfies that $y_1 < y_i < y_j$ with k_i and k_j strictly positive. The gradients of the constraints at the solution are linearly independent and, thus, we can use the KKT conditions of the problem. We denote the dual variables of the equality constraints in (23) with λ, μ and η, respectively. The

KKT conditions for k_i and k_j imply that $f(y) := 1/(1+y)+\lambda y+\mu/y+\eta = 0$ for $y = y_i$ and $y = y_j$. Note that f is continuously differentiable in $(0, 1)$. The mean value theorem implies that there is a point $y_i < \gamma < y_j$ where the derivative $f'(\gamma)$ vanishes. In addition, the KKT conditions for y_i and y_j say that $f'(y_i) = f'(y_j) = 0$. To finish, we prove that it is a contradiction that $f'(\cdot)$ has three zeros. Indeed, $f'(y) = -1/(1 + y)^2 + \lambda - \mu/y^2$ equals zero if and only if $\mu/y^2 + 1/(1 + y)^2 = \lambda$. Regardless of the sign of μ, the function $\mu/y^2 + 1/(1+y)^2$ is unimodal in (0,1). Therefore, it takes any value λ at most twice, which establishes the contradiction. □

The previous result implies that the following reformulation of (24) with values of y equal to y_1 or y_i provides a closed-form upper bound on the price of anarchy. We will see that this upper bound is almost tight for any value of $\sigma \geqslant 2$.

$$\begin{aligned} &\sup \ \frac{\sigma}{4}\,\frac{1-y_1^2}{y_1}\left(2-\frac{2(k_1+1)(y_i-y_1)}{(1+y_1)(1+y_i)}-n\frac{1-y_i}{1+y_i}\right)\\ &\text{s.t. } (k_1+1)(y_i-y_1)+n(1-y_i)=2\\ &\qquad (k_1+1)(y_i-y_1)\frac{1+y_1y_i}{y_1y_i}+n(1-y_i)\frac{1+y_i}{y_i}=\sigma\left(\frac{1}{y_1}-y_1\right)\\ &\qquad 0\leqslant y_1\leqslant y_i\leqslant 1, k_1\geqslant 0, n\geqslant 3. \end{aligned} \tag{25}$$

To solve this problem, we first solve the linear system for k_1 and n given by the equality constraints. Then we plug the result back into the objective function. After some straightforward calculations, this procedure leads us to rewrite (25) as:

$$\max_{0\leqslant y_1\leqslant y_i\leqslant 1} \frac{\sigma}{4}\,\frac{1-y_1}{y_1}\left(\frac{2y_1+y_i(2+4y_1-\sigma(1-y_1^2))}{1+y_i}\right).$$

Observe that the objective is a rational function of y_i, and a simple calculation shows that it is increasing if and only if $y_1 \geqslant 1 - 2/\sigma$. If the rational function were decreasing in an optimal solution, that would imply that $y_i = y_1$, which would evaluate to 1 in the maximum above. Since that cannot be the case, the rational function has to be increasing. Therefore, $y_i = 1$ in an optimal solution, thus making the previous maximum equal to

$$\max_{1-2/\sigma\leqslant y_1\leqslant 1} \frac{\sigma}{4}\,\frac{1-y_1}{y_1}\left(1+3y_1-\frac{\sigma(1-y_1^2)}{2}\right),$$

which is strictly greater than 1. Notice that although Corollary 13 claimed that in an optimal solution $y_i < 1$, our solution satisfies that $y_i = 1$. The reason is that the highest inefficiency at equilibrium is achieved with an infinite number of producers. Therefore, in the limit when $n \to \infty$, y_i equals 1.

To maximize the previous expression, we set its derivative to zero, and find the roots of $-2\sigma y_1^3 + (\sigma - 6)y_1^2 + (\sigma - 2)$. For $\sigma \geqslant 2$, the largest real root $y(\sigma)$ is in the interval $[1 - 2/\sigma, 1]$ (actually, there is only one root when $2 \leqslant \sigma \leqslant 2.33462\ldots$). Solving the cubic equation in closed form, we conclude that the optimal solution for y_1 is

$$y(\sigma) := (v+\sqrt{r})^{1/3}+(v-\sqrt{r})^{1/3}+\frac{\sigma-6}{6\sigma}\,, \tag{26}$$

where

$$v=\frac{55}{216}-\frac{7}{12\sigma}+\frac{1}{2\sigma^2}-\frac{1}{\sigma^3} \qquad \text{and} \qquad r=\left(\frac{55}{216}-\frac{7}{12\sigma}+\frac{1}{2\sigma^2}-\frac{1}{\sigma^3}\right)^2-\left(\frac{1}{6}-\frac{1}{\sigma}\right)^6.$$

Plugging the value back into the objective, we get an upper bound on the price of anarchy valid for all $\sigma \geqslant 2$. We summarize these results in the following theorem.

Theorem 17. *If $\sigma \geqslant 2$, then*

$$POA(\sigma) \leqslant \frac{\sigma}{4}\left(\frac{1-y(\sigma)}{y(\sigma)}\right)\left(1+3y(\sigma)-\frac{\sigma(1-y(\sigma)^2)}{2}\right),$$

where $y(\sigma)$ is given by (26). This bound is tight infinitely often for $\sigma \to \infty$. In particular, the value is exactly $3/2$ when $\sigma = 2$.

Proof. We have already proved the upper bound, we still need to show that the bound is tight for infinitely many values of σ. We specifically show that the bound is tight whenever

$$k_1(\sigma) = \sigma\frac{1+y(\sigma)}{1-y(\sigma)} - \left(\frac{1+y(\sigma)}{1-y(\sigma)}\right)^2 \tag{27}$$

is integral. Indeed, in this case we can evaluate (25) with $k_1 = k_1(\sigma)$, $y_1 = y(\sigma)$, a large enough n, and the appropriate value of y_i (which will be very close to 1). It is not hard to see that the objective value approaches that in the claim of this theorem when $n \to \infty$. Since in this situation k_1 and n are integral, we can construct a sequence of instances whose inefficiency asymptotically equals our bound. Observe that because $y(\sigma) = 1 - 2/\sigma + o(1/\sigma)$, $k_1(\sigma)$ increases to infinity, so that it is integral for infinitely many values of σ. In particular for $\sigma = 2$, $y(\sigma) = 0$ and then $k_1(\sigma) = 1$. □

Notice that the characterization proposed in this section allows us to have significantly improved bounds compared to those in the previous section (which were arbitrarily loose for $\sigma \to 2$). Our new bound is tight in that case, and also when $\sigma \to \infty$. Moreover, we have established that the bound is tight whenever the resulting k_1 is integral, which happens for infinite values of σ. For example, the numerical values of σ that correspond to $k_1(\sigma)$ equal to 1, 2, 3, and 4, are 2, 4.63436, 6.75759, 8.81767, respectively. The resulting lower and upper bounds of the price of anarchy match at 1.5, 1.01858, 1.00747, and 1.00407. Observe that already for values of σ slightly larger than 4, the price of anarchy is below 2%.

We now compute an almost matching lower bound for other values of σ (different from those leading to an integer $k_1(\sigma)$) by restricting solutions to have only two values. A simple way of doing this is to round $k_1(\sigma)$, up and down from the value given by (27), and then evaluate the objective function of (25) for the optimal values of n, y_1 and y_i. Indeed, fixing σ, k_1 and n, one can use the constraints to solve for y_1 and y_i. It turns out that y_1 is the solution of a cubic equation similar to (26). Since we are interested in the limit, we solve the problem for the limit of n growing to infinity.

Proposition 18. *If $\sigma > 2$, then*

$$POA(\sigma) \geqslant \sigma\frac{2(1-y_1^2)-(k_1+1)(1-y_1)^3}{8y_1},$$

where

$$y_1 = \frac{k_1 - 1 + \sqrt{\sigma^2 - 4k_1}}{\sigma + k_1 + 1}.$$

Proof. In the limit when $n \to \infty$, we have that $y_i \to 1$ and $n(1-y_i) \to 2-(k_1+1)(1-y_1)$. After some algebra, the second constraint of (25) yields the value of y_1 in the claim. □

Using the last proposition, we can get a very tight lower bound on the price of anarchy. To this end we evaluate the lower bound given in the proposition for $\lfloor k_1(\sigma)\rfloor$ and $\lceil k_1(\sigma)\rceil$ (as defined in (27)), and take the maximum of the two values. For the special case of $\sigma = 2$, the bound of Proposition 18 is not well defined because $y_1 = 0$. Nevertheless, in this case it is easy to construct the worst case instance. Indeed, it is enough to take $n = 3$ and three different values of y. A calculation shows that, in the limit, this worst-case instance has a price of anarchy of $3/2$, matching the upper bound of Theorem 17.

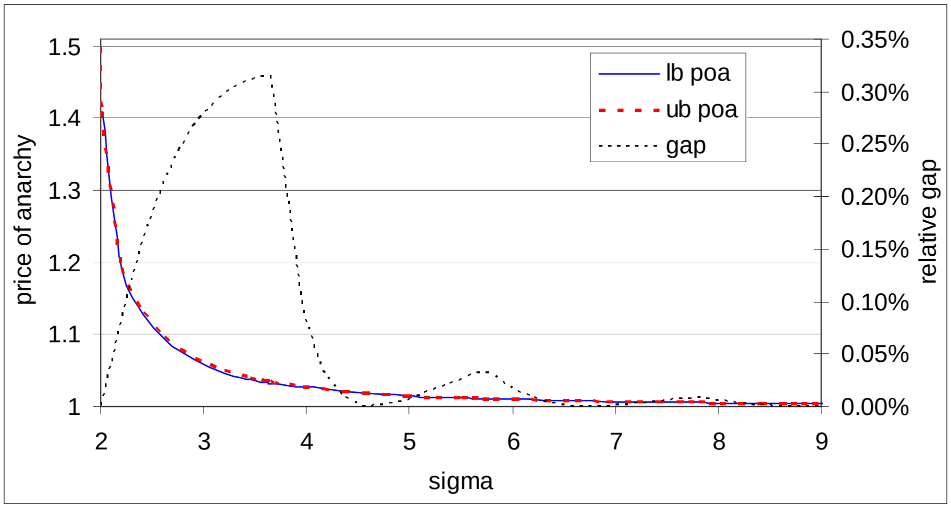

FIGURE 1. Lower bound and upper bounds for the inefficiency of Nash equilibria. The vertical axis on the right displays the relative distance between the two bounds.

Figure 1 depicts both bounds as a function of σ. The figure also includes the relative gap $(ub(\sigma) - lb(\sigma))/ub(\sigma)$ between the two bounds on the secondary vertical axis. Notice how the error goes down to zero every time that $k_1(\sigma)$ approaches an integer. The inflection points in the curve depicting the relative gap correspond to changes in the value of k_1. As an example, the upper bound evaluates to approximately 1.02717 when $\sigma = 4$ while the lower bound evaluates to 1.02642. The gap between the two is approximately 0.074%. From the figure it is clear that the two bounds are very tight. Indeed, the worst relative distance between the lower and upper bounds is 0.316% for a value of $\sigma \approx 3.65$.

6. Future Directions

There are several possible extensions that would be interesting and relevant to explore. First, although we have assumed in this article that products are perfect substitutes, in some situations this may not be the case. More general network structures such as series-parallel networks can be used to model markets with both horizontal and vertical competition. Another interesting direction is to allow for fixed costs and other types of cost functions, as well as more general demand structures such as oligopsonies to model that some consumers may have more market power than others. Both of these more general market structures have been considered for the consumer game when producers are non-strategic and exogenous. Indeed, for supply functions fixed a-priori, a series of papers by Roughgarden and Tardos (2002, 2004), Correa, Schulz, and Stier-Moses (2004, 2008), Perakis (2007), and Cominetti, Correa, and Stier-Moses (2007), among others, look at the inefficiency introduced by competition for increasingly more general cost functions, demand structures, and market structures. Some of the previous references also allow for the presence of side constraints which may be used to model production capacity. If one does the analysis of Section 5 for fixed supply functions, an equilibrium of the second-stage game coincides with a centralized assignment that minimizes the total price paid. This happens because the negative externalities are proportional to the marginal costs (Dafermos and Sparrow 1969). Instead, the assumptions used by the results cited above cannot rule out that equilibria are inefficient, implying that a more general version of our model will have to consider two sources of inefficiency: the markups may distort costs and the consumers may be assigned sub-optimally. Finally, we do not know the extent of the inefficiency of equilibria in the case of elastic demands with general cost functions.

Acknowledgments. This research was partially funded by Instituto Sistemas Complejos de Ingeniería. The authors wish to thank Roberto Cominetti for stimulating discussions, and the participants of seminars at Columbia Business School, at the School of Industrial and Systems Engineering of Georgia Tech, at the Ecole Polytechnique Fédéral de Lausanne, and at the Centro de Economía Aplicada of Universidad de Chile, for their questions, comments and suggestions.

References

Acemoglu, D., K. Bimpikis, and A. Ozdaglar (2009). Price and capacity competition. *Games and Economic Behavior*. Forthcoming.

Acemoglu, D. and A. Ozdaglar (2007). Competition and efficiency in congested markets. *Mathematics of Operations Research 32*(1), 1–31.

Akgün, U. (2004). Mergers with supply function. *Journal of Industrial Economic 52*(4), 535–545.

Baldick, R., R. Grant, and E. Kahn (2004). Theory and application of linear supply function equilibrium in electricity markets. *Journal of Regulatory Economics 25*(2), 143–167.

Cominetti, R., J. R. Correa, and N. E. Stier-Moses (2007). Network games with atomic players. *Operations Research*. Forthcoming.

Correa, J. R., A. S. Schulz, and N. E. Stier-Moses (2004). Selfish routing in capacitated networks. *Mathematics of Operations Research 29*(4), 961–976.

Correa, J. R., A. S. Schulz, and N. E. Stier-Moses (2008). A geometric approach to the price of anarchy in nonatomic congestion games. *Games and Economic Behavior 64*, 457–469.

Dafermos, S. C. and F. T. Sparrow (1969). The traffic assignment problem for a general network. *Journal of Research of the U.S. National Bureau of Standards 73B*, 91–118.

Engel, E., R. Fischer, and A. Galetovic (2004). Toll competition among congested roads. *Topics in Economic Analysis and Policy 4*(1).

Graham, R. L., D. E. Knuth, and O. Patashnik (1994). *Concrete Mathematics: A Foundation for Computer Science* (2nd ed.). Addison-Wesley, Reading, MA.

Johari, R. and J. N. Tsitsiklis (2004). Network resource allocation and a congestion game. *Mathematics of Operations Research 29*(3), 407–435.

Johari, R. and J. N. Tsitsiklis (2008). Parameterized supply function bidding: equilibrium and welfare. Working paper.

Klemperer, P. D. and M. Meyer (1989). Supply function equilibria in oligopoly under uncertainty. *Econometrica 57*(6), 1243–1277.

Koutsoupias, E. and C. H. Papadimitriou (1999). Worst-case equilibria. In C. Meinel and S. Tison (Eds.), *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, Trier, Germany, Volume 1563 of *Lecture Notes in Computer Science*, pp. 404–413. Springer, Heidelberg.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. Oxford University Press, New York, NY.

McAfee and Porter (2009). A theory of bilateral oligopoly. *Economic Inquiry*. Forthcoming.

Nocedal, J. and S. J. Wright (2006). *Numerical Optimization*. Springer, Berlin.

Perakis, G. (2007). The "price of anarchy" under nonlinear and asymmetric costs. *Mathematics of Operations Research 32*(3), 614–628.

Roughgarden, T. and É. Tardos (2002). How bad is selfish routing? *Journal of the ACM 49*, 236–259.

Roughgarden, T. and É. Tardos (2004). Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior 47*(2), 389–403.

Turnbull, S. (1983). Choosing duopoly solutions by consistent conjectures and by uncertainty. *Economics Letters 13*, 253–258.

Weintraub, G. Y., R. Johari, and B. Van Roy (2008). Investment and market structure in industries with congestion. Working Paper.

Wichiensin, M., M. G. H. Bell, and H. Yang (2007). Impact of congestion charging on the transit market: An inter-modal equilibrium model. *Transportation Research Part A 41*, 703–713.

Xiao, F., H. Yang, and D. Han (2007). Competition and efficiency of private toll roads. *Transportation Research Part B 41*, 292–308.