Transaction Costs, Span of Control and Competitive Equilibrium[1]

Tomoo Kikuchi[a], Kazuo Nishimura[b] and John Stachurski[c]

[a] Department of Economics, National University of Singapore

[b] Institute of Economic Research, Kyoto University

[c] Research School of Economics, Australian National University

February 4, 2014

ABSTRACT. This paper embeds the main ideas from Ronald Coase's famous essay on the theory of the firm into a competitive equilibrium setting with an infinite number of identical price taking producers. By formulating the firm problem recursively, we obtain equilibrium prices at every stage of production, and, from these prices, a unique allocation of productive tasks across firms. In this equilibrium, the first order conditions of firms match precisely with Coase's original conjecture on the determinants of firm size. Analysis of the equilibrium yields predictions on prices, the distribution of value added, characteristics of upstream and downstream firms, and division of the value chain.

JEL Classifications: C62, D21

Keywords: Transaction costs, span of control, competitive equilibrium

[1]The authors thank Thibault Fally, Debraj Ray, Tom Sargent, Makoto Yano and Shenghao Zhu for helpful comments and discussions, and Denis Tan, Ngoc Han Huynh Khanh and Alex Olssen for excellent research assistance. We acknowledge financial support from the Grant-in-Aid for Scientific Research, JSPS, (23000001 and 23330063), Singapore Ministry of Education Academic Research Fund R-122-000-140-112 and ARC Discovery Outstanding Researcher Award DP120100321.

Email: `ecstk@nus.edu.sg`, `nishimura@kier.kyoto-u.ac.jp` `john.stachurski@anu.edu.au`

1. INTRODUCTION

Reflecting on a conversation with an ex-Soviet official wishing to know who was in charge of supplying bread to the city of London, Seabright [28] observed that "there was nothing naive about his question, because the answer 'nobody is in charge' is, when one thinks about it, astonishingly hard to believe." Seabright's observation highlights the ability of market forces to coordinate a vast number of specialized activities in an efficient way, without the need for conscious top-down planning.

In his celebrated essay on the nature of the firm, Coase [7] considered the problem of substitution between planners and markets from a striking new direction. First, he observed that even in free market economies with decentralized production, a great deal of top-down planning does in fact take place. However, the planners are not bureaucrats in government departments with input-output tables; instead they are "managers" and "entrepreneurs" who work in entities called "firms." Second, he observed that the planned interactions within firms can, at least in principle, be replaced by market transactions. Given the efficiency of market coordination and the ability to substitute towards it, why, asked Coase, do firms exist at all? What need is there for these "islands of conscious power in the ocean of unconscious cooperation"?[2]

In addressing the existence and size of firms, Coase sought to provide a theory that was both realistic and tractable "by two of the most powerful instruments of economic analysis developed by Marshall: the idea of margin and that of substitution, together giving the idea of substitution at the margin" [7, p. 386]. He argued that firms exist because there are transaction costs associated with using the market, and hence entrepreneurs and managers can sometimes coordinate production at a lower cost within the firm. On the other hand, since firms do not expand without limit, a countervailing force must be present. Coase called this force "diminishing returns to management." The boundary of the firm is then determined by the point at which the cost of organizing another productive task within the firm is equal to the cost of acquiring a similar input or service through the market.

[2]This phrase from Coase's essay is originally due to [25].

Following Coase's analysis, many economists have expanded on and developed the theory of the firm. Researchers have emphasized the effects of imperfect information, incentive and agency problems, incomplete contracts, property rights, decision rights, and the microfoundations of transaction costs. These studies have been highly valuable in building our understanding of the attributes and functions of firms.

In this paper we adopt a relatively high-level perspective on the firm. We first embed the essential features of Coase's verbal analysis into a competitive model with an arbitrary number of firms. We frame the choice problem of firms recursively, which allows us to reduce the model to a single functional equation stated in terms of model primitives. We prove that this equation has a unique and well-behaved solution, and that this solution identifies a decentralized system of production that is well-defined and coordinated by market prices. We then analyze the structure of production, the relationship between upstream and downstream firms, the overall distribution of firms, and the relationship between span of control, transaction costs and the properties of the vertical production chain.

Through this process, we provide quantitative interpretations of Coase's main insights and generate a number of new predictions vis-à-vis the structure of production. For example, we find that downstream firms are larger than upstream firms in terms of value added per dollar of output, and provide a precise relationship in terms of model primitives. We show that higher transaction costs raise final prices, increase firm sizes and reduce the number of active firms, while higher internal coordination costs raise final prices, but decrease firm sizes and increase the number of active firms for a given level of final output. We also show how the model generates a nontrivial, heavy-tailed and right-skewed distribution of firms, despite the fact that all firms are ex-ante identical.

In terms of theoretical contributions, we show that the equilibrium price function is uniquely defined, strictly convex and continuously differentiable for each set of primitives, and that the optimal actions of firms given this price function are unique, well-defined and continuous. We provide an operator that can be used to compute the equilibrium price function, and a class of candidate functions from which convergence is guaranteed. We derive a first order condition that corresponds to the key marginal condition determining firm boundaries stated verbally by Coase (1937), and add an "Euler" equation relating costs of adjacent firms. We also formulate and solve for equilibria in production networks with multiple upstream partners.

Our paper is related to the earlier transaction-costs-in-equilibrium literature, such as Hahn [11], Heller and Shell [13], and Reffett [23], as well as to several familiar models of competitive industry equilibrium, such as Lucas [17], Hopenhayn [14] and Jovanovic [15]. For example, Lucas [17] derives a competitive allocation exhibiting a nontrival size distribution for firms by adopting a notion of "span of control" for managers that closely parallels Coase's idea of diminishing returns to management. Heterogeneity across firms is driven by differences in management technology that exist ex-ante, and the lack of transaction costs means that the competitive equilibrium coincides with the social optimum. By contrast, in our model all producers are ex-ante identical in every way, and face no ideosyncratic shocks. Ex-post heterogeneity is generated purely by the interaction of firms. Moreover, transaction costs distort prices, and are fundamental to the equilibrium allocation.

Our analytical techniques are based on fixed point methods that provide a detailed picture of prices and production. On a mathematical level, the fixed point problem associated with equilibrium prices resembles that of a Bellman equation, although the associated operator fails to be contractive. Bellman operators that fail to be globally contractive have been tackled recently by Rincón-Zapatero and Rodríguez-Palmero [24], Martins-da-Rocha and Vailakis [18], and Kamihigashi [16]. In this paper we adopt a relatively direct approach, which yields existence, uniqueness and global convergence for the particular problem considered here.

The paper is structured as follows: Section 2 describes the model. The equilibrium is analyzed in section 3. Section 4 reviews implications and predictions, and considers some extensions. Section 5 concludes.

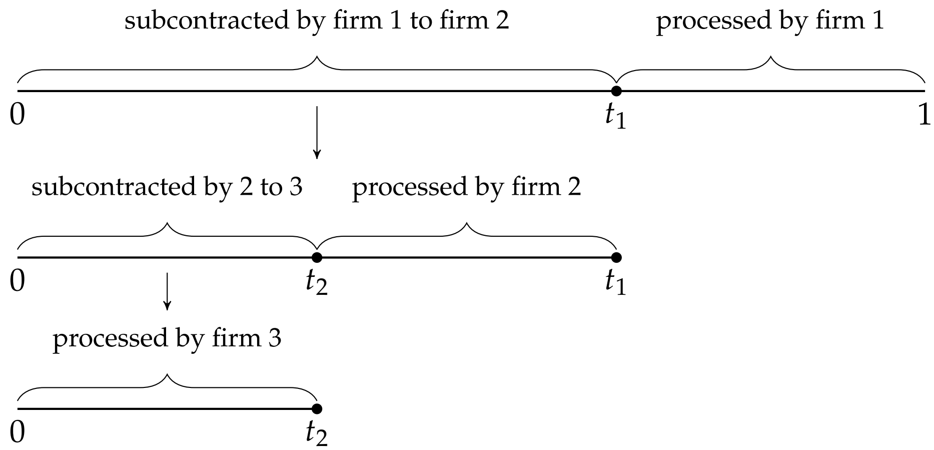


FIGURE 1. Recursive allocation of production tasks

2. THE MODEL

Our focus is supply side and partial equilibrium. For now we will be concerned with the cost of production for a single unit of a single good. At first we will consider a good that is produced through the sequential completion of a number of processing stages. (More complex production structures are treated in section 4.3.) On an intuitive level, we can think of movement from one processing stage to the next as requiring a single specialized task, although this interpretation is not necessary for what follows.

2.1. **The Production Chain.**

In order to provide a sharper marginal analysis, we model the processing stages as a continuum. In particular, the stages are indexed by $t \in [0, 1]$, with $t = 0$ indicating that no tasks have been undertaken and $t = 1$ indicating that the good is complete. Allocation of tasks among firms takes place via subcontracting. The subcontracting scheme is illustrated in figure 1. In this example, an arbitrary firm—henceforth, firm 1—receives a contract to sell one unit of the completed good to a final buyer. Firm 1 then forms a contract with firm 2 to purchase the partially completed good at stage t_1, with the intention of implementing the remaining $1 - t_1$ tasks in-house (i.e., processing from stage t_1 to stage 1). Firm 2 repeats this procedure, forming a contract with firm 3 to purchase the good at stage t_2. In the example in figure 1, firm 3 decides to complete the chain, selecting $t_3 = 0$.

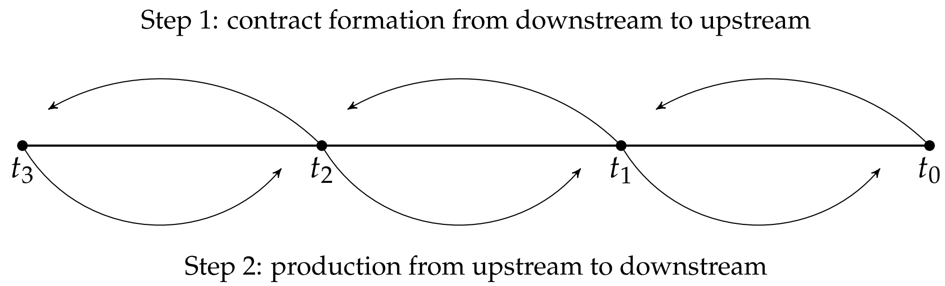


FIGURE 2. Contracts and production

Figure 1 already suggests the recursive nature of the decision problem for each firm. In choosing how many processing stages to subcontract, each successive firm will face essentially the same decision problem as the firm above it in the chain, with the only difference being that the decision space is a subinterval of the decision space for the firm above. Figure 2 emphasizes the fact that contracts evolve backwards from 1 down to 0 (downstream to upstream), and production then unfolds in the opposite direction. In the present example, firm 3 completes processing stages from $t_3 = 0$ up to t_2 and transfers the good to firm 2. Firm 2 then processes from t_2 up to t_1 and transfers the good to firm 1, who processes from t_1 to 1 and delivers the completed good to the final buyer. Figure 3 serves to clarify notation. The length of the interval of stages carried out by firm n is denoted by ℓ_n. In what follows, we typically refer to ℓ_n as the "range" or "number" of tasks carried out by firm n.

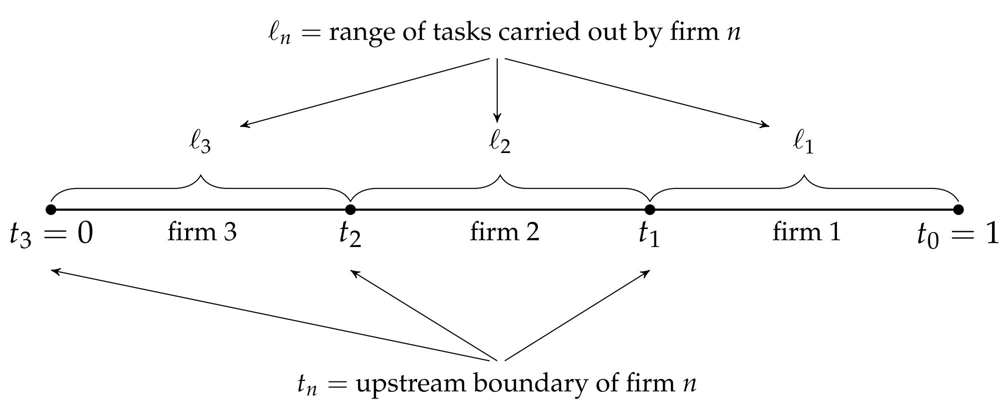


FIGURE 3. Notation

For simplicity the preceding figures have been drawn such that the production chain is equally divided between the firms. As we show below, this is typically not the case in equilibrium.

2.2. **Management and Transaction Costs.** One of the fundamental forces in Coase's theory of the firm is "diminishing returns to management" (Coase, 1937, p. 395), which implies rising costs per task when a firm expands the range of productive activities implemented within its boundaries and coordinated by its managers. Coase argued that rising costs per task were driven by the rapidly expanding informational requirements associated with larger planning problems, leading to high management costs, as well as mistakes, and misallocation of resources. In section 4 we discuss these ideas further. For now we simply take the cost of carrying out ℓ tasks in-house to be $c(\ell)$, and assume that c is strictly convex. Thus, average cost per task rises with the number of tasks performed in-house. We also assume that c is continuously differentiable, with $c(0) = 0$ and $c'(0) > 0$. These assumptions imply that c is strictly increasing.[3]

Diminishing returns to management makes in-house production expensive. Without a countervailing force, the "equilibrium" size of firms might be infinitesimally small. In Coase's analysis, the countervailing force is provided by transaction costs associated with buying and selling through the market. Section 4 gives a detailed discussion of transaction costs. One example is the cost of negotiating, drafting monitoring and enforcing contracts with suppliers. (Suppose for example that contracts are complete but not free.) For now we follow Arrow [1] in simply regarding transaction costs as a wedge between the buyer's and seller's prices. As shown in footnote 6 below, in our model it matters little whether the transaction cost is borne by the buyer, the seller or both. Hence we assume that the cost is borne only by the

[3]The cost function c is assumed to already represent current management best practice, in the sense that no further rearrangement of management structure or internal organization can obtain a lower cost of in-house production. Also note that the cost of carrying out ℓ tasks $\ell = s - t$ depends only on the difference $s - t$ rather than s directly. In other words, all tasks are homogeneous. While extensions might consider other cases, our interest is in equilibrium prices and choices of firms in the base case where tasks are ex-ante identical.

buyer. In particular, when two firms agree to a trade at face value v, the buyer's total outlay is taken to be $\delta v + \epsilon$, where $\delta > 1$. The seller receives only v, and the difference is paid to agents outside the model.

The proportional component δ corresponds to those transaction costs that tend to rise with the face value of the transaction, such as insurance, contracts, sales taxes, trade credit, currency hedging, search costs, bribes and so forth. The second component ϵ corresponds to other transaction costs that are more naturally thought of as constant in face value (transportation is one example). For the sake of simplicity, in all of what follows we take $\epsilon = 0$. This assumption is not entirely innocuous, since additive components will in general also be present, but it does serve to streamline the model.[4]

Throughout the paper, our convention will be that the phrase "transaction costs" refers only to transactions that take place through the market, rather than within the firm.

2.3. **Profit Maximization.** Our remaining assumptions are generic competitive assumptions, with the intention being that equilibrium outcomes will be driven by the ideas described above, rather than additional specific structure. In particular, we assume that all firms are ex-ante identical and act as price takers, contracts are complete and fully enforcible, information is perfect, and active firms are surrounded by an infinite fringe of competitive firms ready to step in on the buyer or seller side should it be profitable to do so.[5] There are no fixed costs or barriers to entry. As a result, no holdup occurs in our model.

Let $p(t)$ represent the price of the good completed up to stage t. To begin the process of determining p, consider an arbitrary firm that enters a contract to sell the good at stage $s \in (0, 1]$, and subsequently decides to purchase the good at stage $t \leq s$ (recall the discussion of contract formulation in section 2.1). The firm undertakes the remaining $\ell = s - t$ tasks in house, and its total costs are given by the sum of its

[4]Another possibility found in the finance literature is that transaction costs are proportional to volume rather than value [19]. We do not pursue this possibility here.

[5]For a recent model of firm boundaries with incomplete information, see Schmitz [27].

processing costs $c(s-t)$ and the gross input cost $\delta p(t)$. Recalling that transaction costs are incurred only by the buyer, profits are $p(s) - c(s-t) - \delta p(t)$. For this firm, s is given and t is a choice variable. If t is chosen to minimize costs, then profits become

$$p(s) - \min_{t \leq s} \{c(s-t) + \delta p(t)\}. \tag{1}$$

Here and below, the restriction $0 \leq t$ in the minimum is understood.

3. Equilibrium

The model can now be closed by zero profit and boundary conditions. This section describes the resulting equilibrium prices and implied production structure. Throughout we assume the conditions on c and δ imposed in section 2.2.

3.1. **Equilibrium Prices.** In order to determine equilibrium prices, we apply the zero profit condition to (1), obtaining $p(s) = \min_{t \leq s} \{c(s-t) + \delta p(t)\}$ for all $s \in (0,1]$. The remaining value $p(0)$ is the revenue of firms that supply the initial inputs to production. We impose a zero profits condition in this sector as well, implying that $p(0)$ is equal to the cost of producing these inputs. To simplify notation, we assume that this cost is zero, and hence the boundary condition is $p(0) = 0$. Putting these restrictions together leads us to the equilibrium price equation

$$p(s) = \min_{t \leq s} \{c(s-t) + \delta p(t)\} \quad \text{for all} \quad s \in [0,1]. \tag{2}$$

This is a functional equation in candidate price functions $p \colon [0,1] \to \mathbb{R}_+$. As is clear from inspection, it imposes the boundary condition $p(0) = 0$ (since $c(0) = 0$ and $\delta > 1$). While trades do not take place at every $s \in [0,1]$ in equilibrium, we show that a solution p of functional equation (2) can be used via a simple recursive procedure to determine the set of active firms and the prices at which they trade.[6]

[6] At this stage we can clarify why it matters little whether we place transaction costs on (a) just the buyer side, or (b) both the buyer and seller. Suppose for example that the seller also faces a cost, receiving only fraction $\gamma < 1$ of any sale. The profit function in section 2.3 then becomes $\pi(s,t) = \gamma p(s) - c(s-t) - \delta p(t)$. Minimizing over $t \leq s$ and setting profits to zero yields $p(s) = \min_{t \leq s}\{c(s-t)\gamma^{-1} + \delta\gamma^{-1} p(t)\}$, which is analogous to (2). Nothing substantial has changed, since $\delta/\gamma > 1$, and since c/γ inherits from c all the properties of the cost function stated in section 2.2.

It is by no means clear that a solution to (2) exists, or if it is unique. To understand the problem, let's convert the functional equation into a fixed point problem by introducing an operator T that maps p to Tp via

$$Tp(s) = \min_{t \le s} \{c(s-t) + \delta p(t)\} \quad \text{for all} \quad s \in [0,1]. \tag{3}$$

Note that any fixed point of T is a solution to (2) and vice versa. In fact, T is analogous to a Bellman operator, with p corresponding to a value function and δ to a discount factor. However, in the standard stationary dynamic programming setting, all of the fundamental theory is driven by contraction mapping arguments, and the contraction property turns on the assumption that the discount factor is strictly less than one. Here, however, the "discount factor" δ is strictly greater than one, and T in (3) is not a contraction in any obvious metric. Indeed, $T^n p$ diverges for many choices of p, even when continuous and bounded. For example, if $p \equiv 1$, then $T^n p = \delta^n 1$, which diverges to $+\infty$.

Despite all of these concerns, it turns out that there exists a restricted domain for T consisting of a natural class of candidate equilibria on which T is extremely well-behaved. As we will see, T has a fixed point in this class, the fixed point is unique, and $T^n p$ converges to this fixed point uniformly for any initial p chosen from the class. In fact the convergence occurs in finite time, and the number of iterations required can be calculated *a priori*. Therefore, computing equilibria is straightforward.

We start our search for a reasonable class of candidate equilibria by observing that any equilibrium price function p should satisfy $p \le c$ pointwise on $[0,1]$. The reason is that a single firm can always implement the entire process up to stage s, at cost $c(s)$. Equilibrium outcomes must do at least as well in terms of cost, implying $p(s) \le c(s)$. This gives an upper bound for prices. Working in the other direction, intuition suggests that an equilibrium function p should satisfy $p(s) \ge sc'(0)$ for all $s \in [0,1]$. To see this, observe that since c is strictly convex, average cost per task $c(\ell)/\ell$ decreases as ℓ gets smaller. If we consider the situation when transaction costs are zero (i.e., $\delta = 1$), this property will encourage firms to enter without limit. If there are n firms involved in producing up to stage s, each producing the small quantity s/n, then, since $\delta = 1$, the aggregate cost of producing to stage s is just $n\, c(s/n)$, and $n\, c(s/n) = s\, c(s/n)/(s/n) \to sc'(0)$ as $n \to \infty$. This suggests that $sc'(0)$ will be a lower bound for costs, and hence, in any equilibrium with $\delta > 1$, we should see $p(s) \geq sc'(0)$.

So far we have proposed an upper and a lower bound for equilibrium prices. In addition it is natural to anticipate that equilibrium prices will be (at least weakly) increasing in processing stage s. Furthermore, given the continuity and convexity of c it seems likely that equilibrium prices will be continuous and convex. Putting these restrictions together, we let $\mathscr{P}$ be the set of convex increasing continuous functions $p \colon [0,1] \to \mathbb{R}$ such that $c'(0)s \leq p(s) \leq c(s)$ for all $0 \leq s \leq 1$, and search for solutions to (2) in $\mathscr{P}$. Our first theorem shows, among other things, that T maps the set $\mathscr{P}$ into itself, confirming our intuition that $\mathscr{P}$ is a natural class of candidate equilibria. In the statement of the theorem, $\bar{s}$ denotes the largest point in $(0,1]$ satisfying $c'(\bar{s}) \leq \delta c'(0)$.

Theorem 3.1. *Under our assumptions the following statements are true:*

1. *The operator T maps $\mathscr{P}$ into itself.*
2. *T has one and only one fixed point in $\mathscr{P}$, and this fixed point p^* is the unique solution in $\mathscr{P}$ to the pricing equation (2).*
3. *For all $p \in \mathscr{P}$ we have $T^n p = p^*$ whenever $n \geq 1/\bar{s}$.*

The proof is deferred to section 6. The theorem tells us that, for each pair of primitives (δ, c), there exists exactly one equilibrium price function p^*. Given our definition of $\mathscr{P}$, it also tell us that p^* is continuous, increasing and convex. In fact we can strengthen these properties without additional assumptions, from increasing to strictly increasing, from convex to strictly convex, and from continuous to continuously differentiable:

Proposition 3.1. *The equilibrium price function p^* is strictly convex, strictly increasing and continuously differentiable on the interior of its domain, with slope less than $c'(1)$.*

Here the strict convexity of p^* is caused not only by strictly diminishing returns to management, but also by positive transaction costs. Intuitively, strict convexity arises because positive transaction costs prevent firms from eliminating diminishing returns to management by infinite subdivision of tasks across firms.

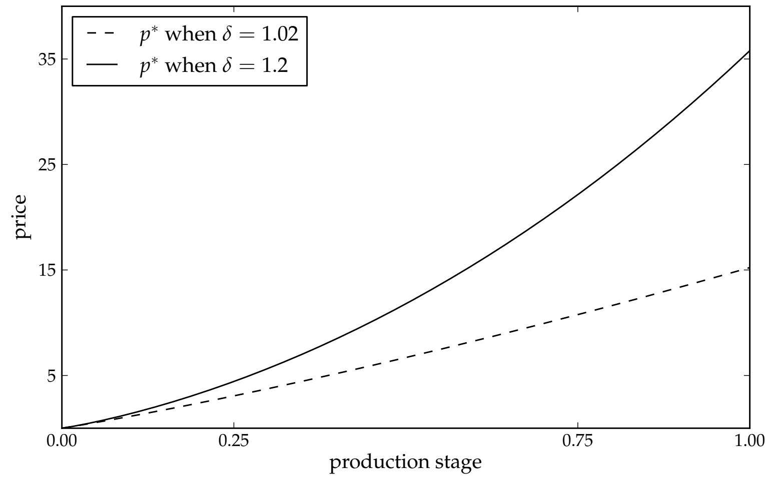


FIGURE 4. Equilibrium price functions when $c(\ell) = e^{\theta\ell} - 1$

Returning to theorem 3.1, part 3 of the theorem is a particularly strong result in terms of computation because it tells us that $T^n p$ converges to p^* in finite time, and, moreover, our finite time bound is uniform over the set of $p \in \mathscr{P}$. However, we emphasize that the convergence result is only valid when the initial condition p is chosen from $\mathscr{P}$. For many other seemingly well-behaved initial conditions, the sequence $T^n p$ will diverge.

Two equilibrium price functions are shown in figure 4, computed by iterating with T from some initial $p \in \mathscr{P}$. In both cases we set $c(\ell) = e^{\theta\ell} - 1$ with $\theta = 10$. The dashed line corresponds to $\delta = 1.02$, while the solid line is for $\delta = 1.2$. Not surprisingly, when transaction costs rise so do prices. (That this relationship always holds is confirmed in proposition 3.4 below.) The figure also shows the strict convexity of p^* obtained in proposition 3.1, and that higher transaction costs cause more curvature in p^* for given c. This is because an increase in transaction costs leads to a smaller number of firms, each doing a larger range of tasks—and hence incurring greater diminishing returns to management.

3.2. **Equilibrium Choices.** In equilibrium, the optimal choices of firms are described by the solution to the cost minimization problem (1) evaluated at the equilibrium

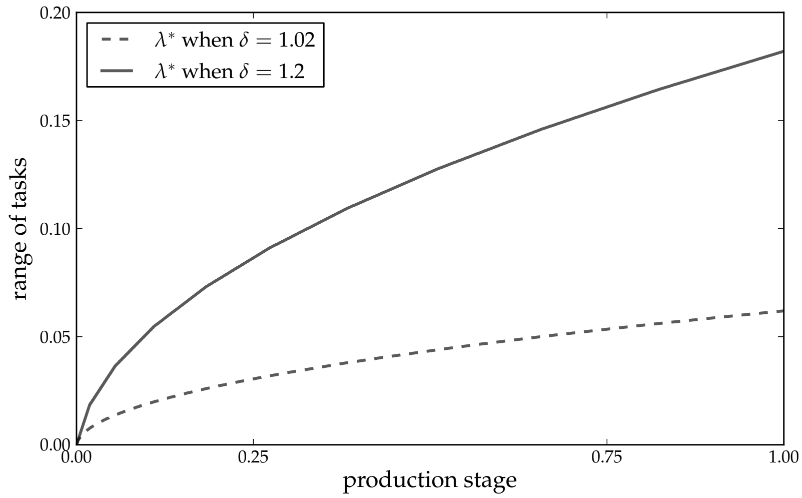
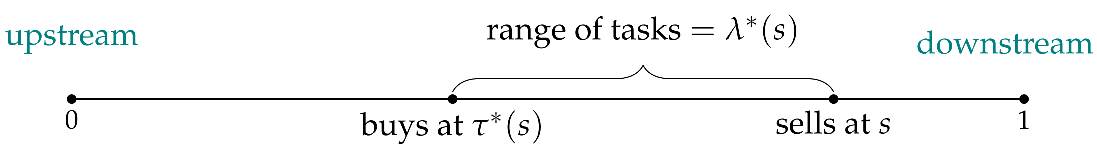


FIGURE 5. Optimal range of in-house tasks as a function of s

price function p^*. As a matter of notation, we set

$$\tau^*(s) := \arg\min_{t \leq s}\{c(s-t) + \delta p^*(t)\} \quad \text{and} \quad \lambda^*(s) := s - \tau^*(s). \tag{4}$$

As shown in the next figure, $\tau^*(s)$ is optimal upstream boundary for a firm that is contracted to deliver the good at stage s and faces equilibrium prices p^*, and $\lambda^*(s)$ is the optimal range of in-house tasks



The function λ^* is plotted for $\delta = 1.2$ and $\delta = 1.02$ in figure 5. Other parameters are the same as for figure 4. We see that $\lambda^*(s)$ increases with production stage s, and that a rise in transaction costs leads to a pointwise increase in $\lambda^*(s)$, so that larger transaction costs means larger firms at all stages of the production chain. We return to both points later on.

In view of the convexity and differentiability results obtained in proposition 3.1, interior solutions to (4) are characterized by the first order condition

$$\delta(p^*)'(\tau^*(s)) = c'(s - \tau^*(s)). \tag{5}$$

Equation (5) states that the upstream boundary of the firm is determined as the processing stage at which the marginal cost of in-house production is equal to the cost of market acquisition. This condition corresponds precisely to the fundamental marginal condition derived verbally by Coase [7] that formed the essence of his theory: "a firm will tend to expand until the costs of organizing an extra transaction within the firm become equal to the costs of carrying out the same transaction by means of an exchange on the open market…" (Coase, 1937, p. 395).

In order to obtain a uniquely defined industrial structure, it is necessary that the functions τ^* and λ^* are well-defined and single-valued. That these properties hold follows from the convexity and continuity of p^* established above. In addition,

Theorem 3.2. *Both τ^* and λ^* are increasing and Lipschitz continuous everywhere on $[0, 1]$. Moreover, for $s \in (0, 1)$, the derivative of p^* satisfies*

$$(p^*)'(s) = c'(\lambda^*(s)). \tag{6}$$

Equation (6) follows from $p^*(s) = \min_{t \leq s} \{c(s - t) + \delta p^*(t)\}$ and the envelope theorem. A full proof can be found in section 6. The monotonicity of τ^* and λ^* is not entirely obvious from our assumptions, and has an interesting interpretation in terms of the differences between upstream and downstream firms. We return to this point below.

3.3. **Allocation of Tasks.** Given p^*, the equilibrium structure of the vertical production chain introduced in figures 1–3 can be determined recursively as follows: Let τ^* be the function defined in (4). Since, by definition, firm 1 sells the completed good at stage $s = 1$, firm 1 buys from firm 2 at stage $\tau^*(1)$. Letting $t_0^* := 1$, and $t_1^* := \tau^*(1)$, we can write this as $t_1^* = \tau^*(t_0^*)$. Since firm 2 sells at t_1^*, it buys at $t_2^* := \tau^*(t_1^*)$. More generally, firm n buys from firm $n + 1$ at

$$t_n^* := \tau^*(t_{n-1}^*) \quad \text{with} \quad t_0^* = 1. \tag{7}$$

This is a difference equation defined by the map τ^*. The iterates $t_1^*, t_2^*, \ldots$ constitute the vertical boundaries of the active firms. Figure 6 shows computed firm boundaries represented by vertical bars for $\delta = 1.05$ and the same cost function as previous figures. Each vertical bar corresponds to a point t_n^*.

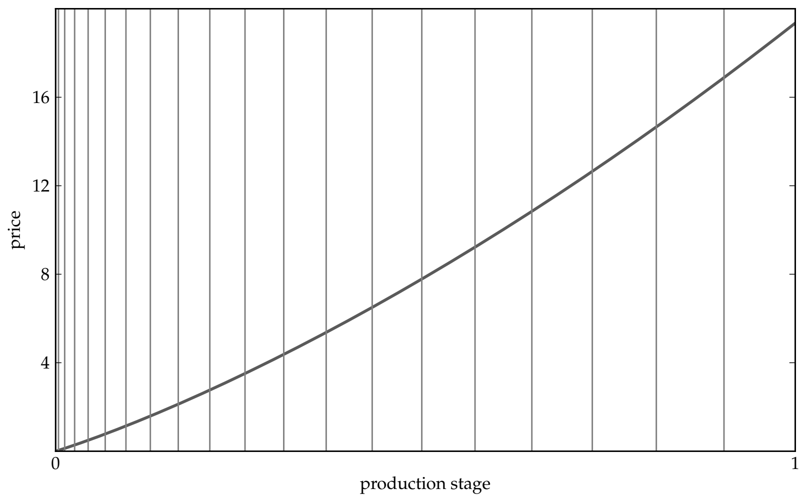


FIGURE 6. Prices and firm boundaries for $\delta = 1.05$

A fundamental question concerning the sequence $\{t_n^*\}$ is whether it attains zero after finitely many iterations. The question is fundamental because production will not actually take place until it does (recall figure 2). It turns out that the answer is affirmative.

Theorem 3.3. *There exists an integer n such that $t_n^* = 0$.*

As a consequence, the equilibrium number of firms corresponding to chosen primitives δ and c is always finite, and uniquely defined as the smallest n such that $t_n^* = 0$. In all of what follows, we denote this integer by N^*.

In-house production (range of tasks) implemented by firm n is given by

$$\ell_n^* := t_{n-1}^* - t_n^*, \qquad n = 1, \ldots, N^*.$$

(See figure 3 on page 6 for clarification.) The sequence $\{\ell_n^*\}$ is the horizontal distance between adjacent vertical bars in figure 8, with n increasing from right to left. We call $\{\ell_n^*\} = \{\ell_1^*, \ldots, \ell_{N^*}^*\}$ the *Coasian allocation of tasks*.

Our model generates an interesting prediction for this allocation that does not appear in Coase's writing (or any subsequent studies):

Proposition 3.2. *The Coasian allocation* $\{\ell_n^*\}$ *satisfies*

$$\delta\, c'(\ell_{n+1}^*) = c'(\ell_n^*). \tag{8}$$

Proposition 3.2 states that marginal in-house cost per task at a given firm is equal to that of its upstream partner multiplied by gross transaction cost. This expression can be thought of as a "Coase-Euler equation," which determines inter-firm efficiency by indicating how two costly forms of coordination (markets and management) are jointly minimized in equilibrium—by equating overall marginal costs between any two adjacent firms.[7]

Since c' is increasing and $\delta > 1$, an immediate consequence of (8) is that $\ell_{n+1}^* \leq \ell_n^*$ for all n. This monotonicity is evident in figure 8. We can obtain the same monotonicity for value added, which is $v_n := p^*(t_{n-1}^*) - p^*(t_n^*)$. Indeed, $\ell_{n+1}^* \leq \ell_n^*$ can also be written as $t_n^* - t_{n+1}^* \leq t_{n-1}^* - t_n^*$. Since p^* is increasing and convex, it follows immediately that $v_{n+1} \leq v_n$. In summary,

Corollary 3.1. *For all n in $1, \ldots, N^* - 1$ we have $\ell_{n+1}^* \leq \ell_n^*$ and $v_{n+1} \leq v_n$.*

The corollary states that the number of in-house tasks and value added both increase the further downstream the firm is in the value chain. The intuition is that the value of the good increases as more processing stages are completed, and with this increase comes a rise in transaction costs (for example, more valuable goods might encourage firms to write more careful—and hence costly—contracts, or to spend more time on search). As a result, downstream firms face higher transaction costs. To economize on these costs, they implement more tasks in-house.[8]

One consequence of the preceding results is that, despite ex-ante identical producers, homogeneous production costs and absence of ideosyncratic shocks, the model generates a nontrivial size distribution of firms. For example, figure 7 shows the

[7]The standard Euler equation connects consumption levels in adjacent time periods in intertemporal consumption models. Our derivation of (8) is essentially identical to the way that the standard Euler equation is derived, by combining first order and envelope conditions.

[8]While it is beyond the scope of the present paper to investigate this prediction empirically, we note that indicies of downstreamness do exist in the empirical literature. See, for example, [8].

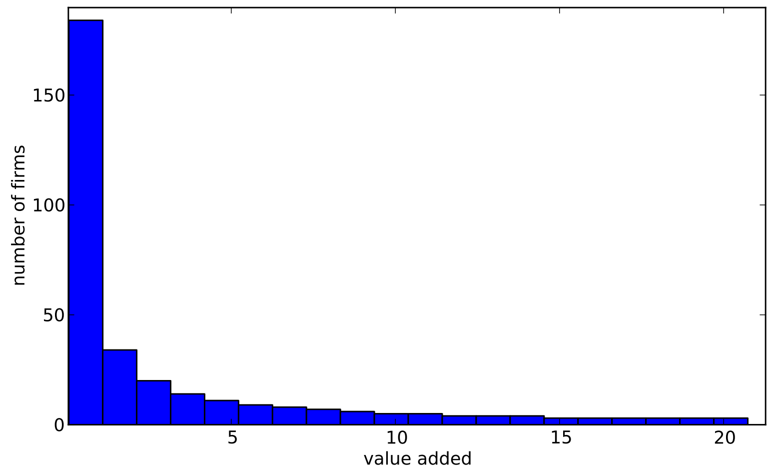


FIGURE 7. Distribution of firms by value added

distribution of firms by value added when $c(\ell) = \exp(\ell^2) - 1$ and $\delta = 1.001$. The distribution is right-skewed and heavy-tailed, both of which are common stylized facts.[9]

3.4. **Output Price.** The preceding discussion all concerns the cost of producing a single unit of the good in question. This cost is given by $p^*(1)$, in the sense that $p^*(1)$ is the amount the most downstream producer must be compensated in order to make zero profits. If, in addition, the production process described above can be replicated any number of times without affecting factor prices, $p^*(1)$ is the long run equilibrium price in the market for the final good.

We can derive an expression for $p^*(1)$ as a weighted sum of the costs incurred in production. The precise relationship is given in the following proposition:

Proposition 3.3. $p^*(1) = \sum_{i=1}^{N^*} \delta^{i-1} c(\ell_i^*)$.

Proof. By definition, we have $p^*(s) = \delta p^*(\tau^*(s)) + c(\lambda^*(s))$ for all s. Evaluating at t_n^* gives the recursion $p^*(t_n^*) = \delta p^*(t_{n+1}^*) + c(\ell_{n+1}^*)$. Combining this with $p^*(t_{N^*}^*) = p^*(0) = 0$ yields the stated expression for $p^*(1)$. $\square$

[9]Of course heavy tailed firm size distributions can be generated by many other phenomena. See, for example, [10].

3.5. **Properties of the Equilibrium.** Let's consider some properties of the equilibrium that we have identified. First, the Coasian allocation $\{\ell_n^*\}$ is not Pareto optimal. The reason is that transaction costs form a wedge between buyer and seller prices, and this distortion of prices twists the allocation away from the social optimum. Put differently, planners can do better than the market in this setting if they have the power to specify allocations directly, without regard to prices. For example, if a planner subdivides the interval $[0,1]$ equally between n firms, then each firm's production cost is $c(1/n)$, and aggregate production cost is $n\,c(1/n)$. Since c is strictly convex, this value is strictly decreasing in n, with infimum given by $c'(0)$. The value $c'(0)$ is strictly less than $p^*(1)$.[10] Hence a social planner can do better than the Coasian equilibrium by dividing production between sufficiently many firms.

There are two important caveats to this discussion. First, the limit $c'(0)$ cannot be attained, as it requires infinitely many firms. It is *because of* the existence of transaction costs that a competitive equilibrium exists in this market. This corresponds to Coase's idea that transaction costs are the friction that gives firms nontrivial size. Second, the comparison between market and planned outcomes provided above should be understood simply as describing a theoretical relationship between the Coasian allocation and classical competitive analysis. It rests entirely on the assumption that the social planner can coordinate production between arbitrarily many firms without cost. This is hardly realistic, especially in our setting where the Coasian allocation greatly improves on what can be done by any single firm.[11] Modeling realistic social planners is far beyond the scope of this paper.

A second important point regarding the equilibrium we have identified is that it is *stable*, in the sense that if any currently inactive firm enters, then, as we show below, at least one firm will make strictly negative profits (and hence decline to participate). However, this opens up the question as to whether equilibrium with zero profits can be recovered by manipulating prices, or by making some more general rearrangement of the firms. Our next theorem shows that this is impossible: for *any*

[10] This follows from the fact that $p^*(s) \geq c'(0)s$ (because $p^* \in \mathscr{P}$) plus strict convexity of p^*.

[11] To give some indication, if transaction costs are 2%, say, and other parameters are as in figure 6, then the production cost $c(1)$ for a single firm is about three orders of magnitude larger than $p^*(1)$.

valid price function and *any* allocation of tasks distinct from $\{\ell_n^*\}$, at least one firm will make strictly negative profits.

To state the theorem, let us call $\boldsymbol{\ell} := (\ell_1, \ldots, \ell_J)$ a feasible allocation of tasks if its elements are nonnegative and $\sum_{n=1}^J \ell_n = 1$. Let $p\colon [0,1] \to \mathbb{R}_+$ be any price function, and let $\boldsymbol{\ell} \in \mathbb{R}^J$ be any feasible allocation of tasks across J firms, where $J \in \mathbb{N}$ is arbitrary. Let π_j be the profits of the j-th firm given the allocation $\boldsymbol{\ell}$ and the price function p.

Theorem 3.4. *If $\boldsymbol{\ell} \neq \boldsymbol{\ell}^*$, then either $p(1) > p^*(1)$ or $\exists\, j \in \{1, \ldots, J\}$ such that $\pi_j < 0$.*

In other words, either (*i*) the alternative production chain will be undercut by the Coasian production chain, which charges the final buyer $p^*(1)$ rather than $p(1)$, or (*ii*) at least one firm in the alternative production chain will make strictly negative profits. In either case the pair $(p, \boldsymbol{\ell})$ is not an equilibrium.

3.6. **Comparative Statics.** In our model, variations in transaction costs shift equilibrium outcomes monotonically in the directions intuition suggests. A rise in transaction costs causes prices to rise, the size of firms to increase, and the equilibrium number of firms to fall. The next proposition gives details.

Proposition 3.4. *If $\delta_a \leq \delta_b$, then the following inequalities hold:*

(1) $p_a^* \leq p_b^*$ *on* $[0,1]$
(2) $\ell_n^a < \ell_n^b$ *for all active firms, and* $N_b^* \leq N_a^*$

Here p_a^* be the equilibrium price function for transaction cost δ_a, p_b^* is that for δ_b, and so on. These outcomes can be observed by comparing the upper and lower panels in figure 8.

4. Discussion

This section reflects in more detail on the fundamental primitives that drive the model, and reviews the model's implications and predictions.

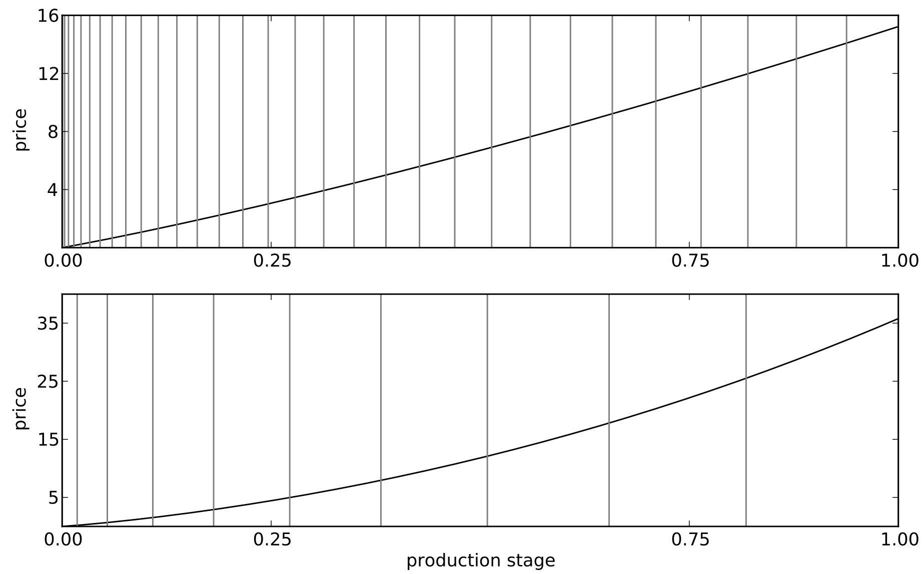


FIGURE 8. Prices and firm boundaries for $\delta = 1.02$ (top) and $\delta = 1.2$ (bottom)

4.1. **Transaction Costs.** On a theoretical level, recognition of the systemic impact of transaction costs dates back to Adam Smith ([29]). Smith famously argued that the division of labor is limited by the extent of the market. Pushing the analysis one step further, he also observed that the extent of the market is itself limited by transportation costs ([29], p. 31). In other words, transaction costs play a fundamental role in limiting specialization (see also [32, 2]). Other commonly cited transaction costs include transaction fees; taxes, bribes and theft associated with transactions; search, bargaining and information costs; trade through middlemen; costs of assessing credit worthiness and reliability; and the costs associated with negotiating, writing, monitoring and enforcing contracts [7, 30, 31, 22, 4].

The predictions of our model vis-a-vis transaction costs are broadly consistent with empirical and case studies in the literature. For example, in [26] it is noted that the substantial market reforms that occurred in Ghana in the 1980s were followed by a significant fall in average firm size as measured by employment, from 19 to 9 in the fifteen years from 1987. The reforms (market deregulation, enforcement of contracts and stricter penalties for bribe extraction) all suggest a reduction in market transaction costs. While Sandefur presented this contraction in firm sizes as a puzzle, in our model it is a prediction (see proposition 3.4).

The model also serves to illustrate that even if transaction costs are small, they can still have first order effects in terms of determining the structure and overall cost of production because they cascade through the production chain. For example, if $c(\ell) = \exp(\ell^2) - 1$ and $\delta = 1.01$, then a 1% rise in δ leads to a 40% fall in the number of firms and a 90% increase in the price of the final good. The first order effect of transaction costs can also be seen in the exponential terms in the expression $p^*(1) = \sum_{i=1}^{N^*} \delta^{i-1} c(\ell_i)$ for the final price proved in proposition 3.3, and in the comparison of production networks shown in figure 8.[12]

4.2. **Diminishing Returns to Management.** The other fundamental force in Coase's theory of the firm is diminishing returns to management. For Coase, these diminishing returns were driven by the huge informational requirements associated with large planning problems, leading to "mistakes" and misallocation of resources. The difficulty of coordinating production through top down planning was also emphasized by Hayek [12], who highlighted the difficulty of utilizing knowledge not held in its totality by any one individual. Later authors have highlighted additional causes of high average management costs in large firms, such as free-riding, shirking and other incentive problems. See, for example, [9, 21].

In our model, the rate at which internal management costs grow with the number of tasks implemented by the firm depends on the curvature of c. If we let c have the exponential form $c(\ell) = e^{\theta \ell} - 1$, then $\theta = c''(\ell)/c'(\ell)$, so θ parameterizes curvature of c, and hence the intensity of diminishing returns to management. The effect of θ can be obtained as follows. From the Coase-Euler equation (8) we have $\ell_{n+1} = \ell_n - \ln \delta / \theta$. Using this equation, the constraint $\sum_{n=1}^{N^*} \ell_n = 1$ and some algebra, it can

[12] It is important to note that the expression $p^*(1) = \sum_{i=1}^{N^*} \delta^{i-1} c(\ell_i)$ does not in fact imply that prices are exponential in δ. The reason is that when δ changes, firms re-optimize, leading to a change in the sequence $\{\ell_i\}$. In particular, higher transaction costs are mitigated by a reduction in the number of firms.

be shown (cf., lemma 6.8) that the equilibrium number of firms is

$$N^* = \text{int}\left(\frac{1+\sqrt{1+8\theta/\ln\delta}}{2}\right), \tag{9}$$

where $\text{int}(a)$ is the largest integer less than or equal to a. Notice that N^* is increasing in θ. The reason is that more intense diminishing returns to management encourages greater use of the market, and hence smaller firms. Since the number of tasks does not change, smaller firms imply more firms. Conversely, a decrease in curvature leads to larger firms, and, in aggregate, fewer firms and fewer transactions.

This discussion provides quantitative structure to predictions originally made by Coase [7, p. 396] concerning the relationship between diminishing returns to management and the size of firms. Such predictions are hard to confirm empirically, since many innovations that change internal coordination costs (e.g., computer networks) also tend to affect transaction costs (through lower search costs and so on). On the other hand, some major management innovations have reduced the cost of coordinating specialists without affecting external transaction costs. Examples include moving assembly lines and the multi-divisional corporate structure, both of which have aided the growth of larger firms [5].

4.3. **Multiple Trade Partners.** So far we have assumed that production is linearly sequential, and that firms contract with only one supplier. In reality, most vertically integrated firms have multiple upstream partners. Our model generalizes naturally to this case. As an example, consider the tree in figure 9. As before, firm n chooses an interval ℓ_n of tasks to perform in-house, and subcontracts the remainder. In this case, however, the remainder is divided between *two* upstream partners. If a firm contracts to supply the good at stage s, chooses a quantity $\ell \leq s$ to produce in-house, and then divides the remainder $s - \ell$ equally across the two upstream partners, then profits $\pi(s, \ell)$ will be given by revenue $p(s)$ minus total input costs $2\,\delta\, p((s-\ell)/2)$ minus in-house production costs $c(\ell)$. That is,

$$\pi(s,\ell) = p(s) - \delta\, 2\, p((s-\ell)/2) - c(\ell).$$

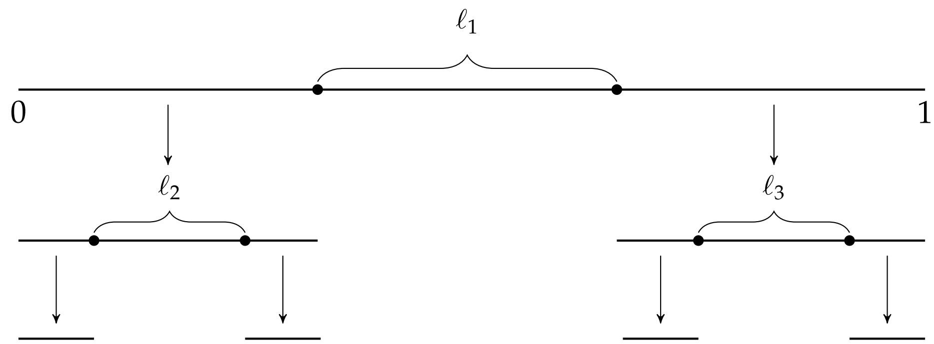


FIGURE 9. Subcontracting with two upstream suppliers

Setting profits to zero, letting $t := s - \ell$ and minimizing with respect to t yields the functional equation

$$p(s) = \min_{t \leq s} \{\delta \, 2 \, p(t/2) + c(s - t)\}.$$

This equation is an immediate generalization of (2). Similar techniques can be employed to show the existence and uniqueness of a solution p^*, to compute p^*, and to compute from p^* the optimal choices of firms and the resulting distribution of firms.

Figure 10 shows the result of these computations when the cost function has the same exponential shape as before and $\delta = 1.02$. Figure 11 shows the result for $\delta = 1.2$. In these graphs, each node represents a firm, and the size of the node is proportional to the value added of that firm. The uppermost node is the final producer. Note that firms towards the top of the network have greater value added, consistent with corollary 3.1. While these production networks are simple symmetric trees, addition of heterogeneity or stochastic choice could potentially generate more realistic features.[13]

5. CONCLUSION

In this paper we embed the key ideas from Coase's famous essay on the theory of the firm into a competitive equilibrium setting with an infinite number of identical price-taking producers. By developing an approach to the equilibrium problem

[13]Competitive trading on networks has been studied in depth by other authors, including [4, 6].

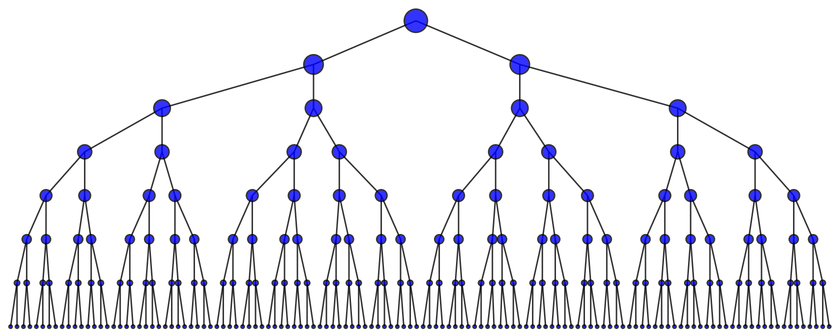


FIGURE 10. Firm network with $\delta = 1.02$ and $c(\ell) = e^{\theta \ell} - 1$

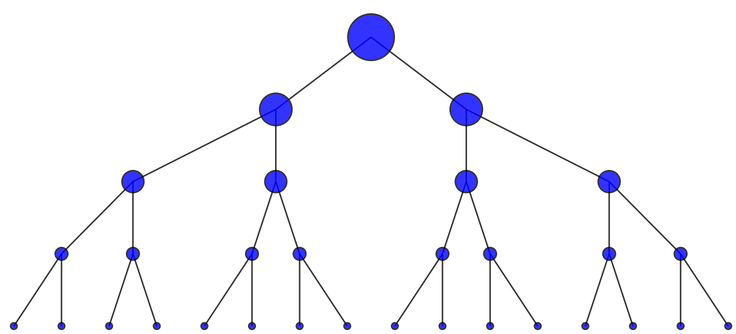


FIGURE 11. Firm network with $\delta = 1.2$ and $c(\ell) = e^{\theta \ell} - 1$

based on recursive subdivision of tasks, we reduce the model to a single functional equation and prove existence and uniqueness of the solution. We then show how the solution uniquely determines prices, actions of firms and vertical division of the value chain.

Through this process, we recover Coase's key insights on the boundaries of individual firms, but now in the form of first order conditions that hold in equilibrium. More importantly, the model allows us to study the implications of Coase's ideas for the structure of production along the value chain, as determined by the interactions between the choices of individual firms and the equilibrium set of prices. These interactions result in a set of new predictions about relative marginal production costs in contracting firms, the relationship between upstream and downstream firms, the impact of transaction costs and the distribution of production.

Our work opens a number of avenues for future research. The techniques for characterizing and computing industry equilibrium are likely to have applications in other fields, such as international trade, or production with failure probabilities and other frictions. In addition, the model presented above is a baseline model in all dimensions, with perfect competition, perfect information, identical firms and identical tasks, and these assumptions can be weakened. The effect of altering contract structures could also be investigated, as could the various possibilities for determining upstream partners in section 4.3. Finally, the model has a number of interesting empirical implications, such as the Coase-Euler equation, and the positive relationship between downstreamness and value added discussed above.

6. Appendix: Remaining Proofs

Our first aim is to prove theorem 3.1. The existence component of the theorem can be established using one of several well-known order-theoretic or topological fixed point results (see, e.g., [20]). Here we pursue a more direct proof, which simultaneously yields existence, uniqueness and a means of computation. We start with a number of preliminary results.

Lemma 6.1. *If $p \in \mathscr{P}$, then Tp is strictly convex.*

Proof. Pick any $0 \leq s_1 < s_2 \leq 1$ and any $\alpha \in (0,1)$. Let $t_i := \arg\min_{t \leq s_i} \{\delta p(t) + c(s_i - t)\}$ for $i = 1, 2$, and $t_3 := \alpha t_1 + (1-\alpha) t_2$. It is easy to check that $0 \leq t_3 \leq \alpha s_1 + (1-\alpha) s_2$, and hence

$$Tp(\alpha s_1 + (1-\alpha) s_2) \leq \delta p(t_3) + c(\alpha s_1 + (1-\alpha) s_2 - t_3).$$

The right-hand side expands out to

$$\delta p[\alpha t_1 + (1-\alpha) t_2] + c[\alpha s_1 - \alpha t_1 + (1-\alpha) s_2 - (1-\alpha) t_2].$$

Using convexity of p and strict convexity of c, we obtain $Tp(\alpha s_1 + (1-\alpha) s_2) < \alpha Tp(s_1) + (1-\alpha) Tp(s_2)$, which is strict convexity Tp. $\square$

Lemma 6.2. *Let $p \in \mathscr{P}$ and let t_p and ℓ_p be the optimal responses, defined by*

$$t_p(s) := \arg\min_{t \le s}\{\delta p(t) + c(s-t)\} \quad \textit{and} \quad \ell_p(s) := s - t_p(s). \tag{10}$$

If s_1 and s_2 are any two points with $0 < s_1 \le s_2 \le 1$, then

(1) *both $t_p(s_1)$ and $\ell_p(s_1)$ are well defined and single-valued.*
(2) *$t_p(s_1) \le t_p(s_2)$ and $t_p(s_2) - t_p(s_1) \le s_2 - s_1$.*
(3) *$\ell_p(s_1) \le \ell_p(s_2)$ and $\ell_p(s_2) - \ell_p(s_1) \le s_2 - s_1$.*

Proof. Since $t \mapsto \delta p(t) + c(s_1 - t)$ is continuous and strictly convex (by convexity of p and strict convexity of c), and since $[0, s_1]$ is compact, existence and uniqueness of $t_p(s_1)$ and $\ell_p(s_1)$ must hold. Regarding the claim that $t_p(s_1) \le t_p(s_2)$, let $t_i := t_p(s_i)$. Suppose instead that $t_1 > t_2$. We aim to show that, in this case,

$$\delta p(t_1) + c(s_2 - t_1) < \delta p(t_2) + c(s_2 - t_2), \tag{11}$$

which contradicts the definition of t_2.[14] To establish (11), observe that t_1 is optimal at s_1 and $t_2 < t_1$, so

$$\delta p(t_1) + c(s_1 - t_1) < \delta p(t_2) + c(s_1 - t_2).$$

$$\therefore \quad \delta p(t_1) + c(s_2 - t_1) < \delta p(t_2) + c(s_1 - t_2) + c(s_2 - t_1) - c(s_1 - t_1).$$

Given that c is strictly convex and $t_2 < t_1$, we have

$$c(s_2 - t_1) - c(s_1 - t_1) < c(s_2 - t_2) - c(s_1 - t_2).$$

Combining this with the last inequality yields (11).

Next we show that $\ell_1 \le \ell_2$, where $\ell_1 := \ell_p(s_1)$ and $\ell_2 := \ell_p(s_2)$. In other words, $\ell_i = \arg\min_{\ell \le s_i}\{\delta p(s_i - \ell) + c(\ell)\}$. The argument is similar to that for t_p, but this time using convexity of p instead of c. To induce the contradiction, we suppose that $\ell_2 < \ell_1$. As a result, we have $0 \le \ell_2 < \ell_1 \le s_1$, and hence ℓ_2 was available when ℓ_1 was chosen. Therefore,

$$\delta p(s_1 - \ell_1) + c(\ell_1) < \delta p(s_1 - \ell_2) + c(\ell_2),$$

[14] Note that $t_1 < s_1 \le s_2$, so t_1 is available when t_2 is chosen.

where the strict inequality is due to the fact that minimizers are unique. Rearranging and adding $\delta p(s_2 - \ell_1)$ to both sides gives

$$\delta p(s_2 - \ell_1) + c(\ell_1) < \delta p(s_2 - \ell_1) - \delta p(s_1 - \ell_1) + \delta p(s_1 - \ell_2) + c(\ell_2).$$

Given that p is convex and $\ell_2 < \ell_1$, we have

$$p(s_2 - \ell_1) - p(s_1 - \ell_1) \leq p(s_2 - \ell_2) - p(s_1 - \ell_2).$$

Combining this with the last inequality, we obtain

$$\delta p(s_2 - \ell_1) + c(\ell_1) < \delta p(s_2 - \ell_2) + c(\ell_2),$$

contradicting optimality of ℓ_2.[15]

To complete the proof of lemma 6.2, we also need to show that $t_p(s_2) - t_p(s_1) \leq s_2 - s_1$, and similarly for ℓ. Starting with the first case, we have

$$t_p(s_2) - t_p(s_1) = s_2 - \ell_p(s_2) - s_1 + \ell_p(s_1) = s_2 - s_1 + \ell_p(s_1) - \ell_p(s_2).$$

As shown above, $\ell_p(s_1) \leq \ell_p(s_2)$, so $t_p(s_2) - t_p(s_1) \leq s_2 - s_1$, as was to be shown. Finally, the corresponding proof for ℓ_p is obtained in the same way, by reversing the roles of t_p and ℓ_p. This concludes the proof of lemma 6.2. □

Recall the constant $\bar{s}$ defined above theorem 3.1, existence of which follows from the conditions in section 2.2 and the intermediate value theorem. Regarding $\bar{s}$ we have the following lemma, which states that the best action for a firm subcontracting at $s \leq \bar{s}$ is to start from stage $t = 0$.

Lemma 6.3. *If $p \in \mathscr{P}$, then $s \leq \bar{s}$ if and only if $\min_{t \leq s} \{\delta p(t) + c(s-t)\} = c(s)$.*

Proof. First suppose that $s \leq \bar{s}$. Seeking a contradiction, suppose there exists a $t \in (0, s]$ such that $\delta p(t) + c(s-t) < c(s)$. Since $p \in \mathscr{P}$ we have $p(t) \geq c'(0)t$ and hence $\delta p(t) \geq \delta c'(0)t = c'(\bar{s})t$. Since $s \leq \bar{s}$, this implies that $\delta p(t) \geq c'(s)t$. Combining these inequalities gives $c'(s)t + c(s-t) < c(s)$, contradicting convexity of c.

[15]Note that $0 \leq \ell_1 \leq s_1 \leq s_2$, so ℓ_1 is available when ℓ_2 is chosen.

Suppose on the other hand that $\inf_{t \leq s} \{\delta p(t) + c(s-t)\} = c(s)$. We claim that $s \leq \bar{s}$, or, equivalently $c'(s) \leq \delta c'(0)$. To see that this is so, observe that since $p \in \mathscr{P}$ we have $p(t) \leq c(t)$, and hence

$$c(s) \leq \{\delta p(t) + c(s-t)\} \leq \{\delta c(t) + c(s-t)\}, \qquad \forall\, t \leq s.$$
$$\therefore \quad \frac{c(s) - c(s-t)}{t} \leq \frac{\delta c(t)}{t} \qquad \forall\, t \leq s.$$
$$\therefore \quad c'(s) \leq \delta c'(0).$$
□

Lemma 6.4. *Let $p \in \mathscr{P}$ and let ℓ_p be as in (10). If $s \geq \bar{s}$, then $\ell_p(s) \geq \bar{s}$. If $s > 0$, then $\ell_p(s) > 0$.*

Proof. By lemma 6.2, ℓ_p is increasing, and hence if $\bar{s} \leq s \leq 1$, then $\ell_p(s) \geq \ell_p(\bar{s}) = \bar{s} - t_p(\bar{s}) = \bar{s}$. By lemma 6.3, if $0 < s \leq \bar{s}$, then $\ell_p(s) = s - t_p(s) = s > 0$. □

Lemma 6.5. *If $p \in \mathscr{P}$, then Tp is differentiable on $(0,1)$ with $(Tp)' = c' \circ \ell_p$.*

Proof. Fix $p \in \mathscr{P}$ and let t_p be as in (10). Fix $s_0 \in (0,1)$. By [3], to show that Tp is differentiable at s_0 it suffices to exhibit an open neighborhood $U \ni s_0$ and a function $w\colon U \to \mathbb{R}$ such that w is convex, differentiable, satisfies $w(s_0) = Tp(s_0)$ and dominates Tp on U. To exhibit such a function, observe that in view of lemma 6.4, we have $t_p(s_0) < s_0$. Now choose an open neighborhood U of s_0 such that $t_p(s_0) < s$ for every $s \in U$. On U, define

$$w(s) := \delta p(t_p(s_0)) + c(s - t_p(s_0)).$$

Clearly w is convex and differentiable on U, and satisfies $w(s_0) = Tp(s_0)$. To see that $w(s) \geq Tp(s)$ when $s \in U$, observe that if $s \in U$ then $0 \leq t_p(s_0) \leq s$, and

$$Tp(s) = \min_{t \leq s}\{\delta p(t) + c(s-t)\} \leq \delta p(t_p(s_0)) + c(s - t_p(s_0)) = w(s).$$

As a result, Tp is differentiable at s_0 with $(Tp)'(s_0) = w'(s_0) = c'(\ell_p(s_0))$. □

Lemma 6.6. *The operator T defined in (3) maps $\mathscr{P}$ into itself.*

Proof. Let p be an arbitrary element of $\mathscr{P}$. To see that $Tp(s) \leq c(s)$ for all $s \in [0,1]$, fix $s \in [0,1]$ and observe that, since $p \in \mathscr{P}$ implies $p(0) = 0$, the definition of T implies $Tp(s) \leq \delta p(0) + c(s+0) = c(s)$. Next we check that $Tp(s) \geq c'(0)s$

for all $s \in [0,1]$. Picking any such s and using the assumption that $p \in \mathscr{P}$, we have $Tp(s) \geq \inf_{t \leq s}\{\delta c'(0)t + c(s-t)\}$. By $\delta > 1$ and convexity of c, we have $\delta c'(0)t + c(s-t) \geq c'(0)t + c(s-t) \geq c'(0)t + c'(0)(s-t) = c'(0)s$. Therefore $Tp(s) \geq \inf_{t \leq s} c'(0)s = c'(0)s$.

It remains to show that Tp is continuous, convex and monotone increasing. That Tp is convex was shown in lemma 6.1. Regarding the other two properties, let ℓ_p and t_p be as defined in (10). By the results in lemma 6.2, these functions are increasing and (Lipschitz) continuous on $[0,1]$. Since $Tp(s) = \delta p(t_p(s)) + c(\ell_p(s))$, it follows that Tp is also increasing and continuous. □

Lemma 6.7. *If $p, q \in \mathscr{P}$, then $T^n p = T^n q$ whenever $n \geq 1/\bar{s}$.*

Proof. The proof is by induction. First we argue that $T^1 p = T^1 q$ on the interval $[0, \bar{s}]$. Next we show that if $T^k p = T^k q$ on $[0, k\bar{s}]$, then $T^{k+1} p = T^{k+1} q$ on $[0, (k+1)\bar{s}]$. Together these two facts imply the claim in lemma 6.7.

To see that $T^1 p = T^1 q$ on $[0, \bar{s}]$, pick any $s \in [0, \bar{s}]$ and recall from lemma 6.3 that if $h \in \mathscr{P}$ and $s \leq \bar{s}$, then $Th(s) = c(s)$. Applying this result to both p and q gives $Tp(s) = Tq(s) = c(s)$. Hence $T^1 p = T^1 q$ on $[0, \bar{s}]$ as claimed.

Turning to the induction step, suppose now that $T^k p = T^k q$ on $[0, k\bar{s}]$, and pick any $s \in [0, (k+1)\bar{s}]$. Let $h \in \mathscr{P}$ be arbitrary, let $\ell_h(s) := \arg\min_{t \leq s}\{\delta h(t) + c(s-t)\}$ and let $t_h(s) := s - \ell_h(s)$. By lemma 6.4, we have $\ell_h(s) \geq \bar{s}$, and hence

$$t_h(s) \leq s - \bar{s} \leq (k+1)\bar{s} - \bar{s} \leq k\bar{s}.$$

In other words, given arbitrary $h \in \mathscr{P}$, the optimal choice at s is less than $k\bar{s}$. Since this is true for $h = T^k p$, we have

$$T^{k+1} p(s) = \min_{t \leq s}\{c(s-t) + \delta T^k p(t)\} = \min_{t \leq k\bar{s}}\{c(s-t) + \delta T^k p(t)\}.$$

Using the induction hypothesis and the preceding argument for $h = T^k q$, this is equal to

$$\min_{t \leq k\bar{s}}\{c(s-t) + \delta T^k q(t)\} = \min_{t \leq s}\{c(s-t) + \delta T^k q(t)\} = T^{k+1} q(s).$$

We have now shown that $T^{k+1} p = T^{k+1} q$ on $[0, (k+1)\bar{s}]$. The proof is complete. □

Proposition 6.1. *The operator T has one and only one fixed point in $\mathscr{P}$.*

Proof. To show existence, let $n \geq 1/\bar{s}$ and fix any $p \in \mathscr{P}$. In view of lemma 6.7, we have $T^n(Tp) = T^n p$. Equivalently, $T(T^n p) = T^n p$. In other words, $T^n p$ is a fixed point of T. Regarding uniqueness, let p and q be two fixed points of T in $\mathscr{P}$, and let $n \geq 1/\bar{s}$. In view of lemma 6.7, we have $p = T^n p = T^n q = q$. □

We can now turn to proving our main results.

Proof of theorem 3.1. That T maps $\mathscr{P}$ to itself was shown in lemma 6.6. Existence and uniqueness of a fixed point p^*, and hence of a solution to the price function (2), follows from proposition 6.1. The claim that $T^n p = p^*$ for all $n \geq 1/\bar{s}$ is immediate from lemma 6.7. □

Proof of proposition 3.1. Since $p^* \in \mathscr{P}$ and the image under T of any function in $\mathscr{P}$ is strictly convex (lemma 6.1), we see that $p^* = Tp^*$ is strictly convex. Letting $p = p^*$ in lemma 6.5, we see that p^* is differentiable, with $(p^*)'(s) = c'(\lambda^*(s))$. Since c and λ^* are continuous, the latter by lemma 6.2, and since $c'(s) > 0$ for all s, the equation $(p^*)'(s) = c'(\lambda^*(s))$ implies that p^* also continuously differentiable and strictly increasing. □

Proof of proposition 3.2. Combining the first order condition in (5) with the envelope condition (6) gives $\delta c'(\lambda^*(\tau^*(s))) = c'(\lambda^*(s))$. Evaluating at $s = t_{n-1}$ completes the proof. □

Proof of theorem 3.2. We have already shown that $(p^*)'(s) = c'(\lambda^*(s))$. The claimed properties on τ^* and λ^* are immediate from lemma 6.2. □

Proof of theorem 3.3. In view of lemma 6.3 we have $\lambda^*(s) = s$ and hence $\tau^*(s) = 0$ whenever $s \leq \bar{s}$, so it suffices to prove that $t_n \leq \bar{s}$ for some n. This must be the case because λ^* is increasing, and hence the amount of in-house production by a firm contracting at $s \geq \bar{s}$ satisfies $\lambda^*(s) \geq \lambda^*(\bar{s}) = \bar{s} > 0$. In other words, for firms contracting above $\bar{s}$, each takes a step of length at least $\bar{s}$. In particular, $t_n \leq 1 - n\bar{s}$. □

Proof of theorem 3.4. Suppose to the contrary that there exists a price function $p\colon [0,1] \to \mathbb{R}_+$ and an allocation $\boldsymbol{\ell} \in \mathbb{R}^J$ such that $p(1) \leq p^*(1)$, $\boldsymbol{\ell} \neq \boldsymbol{\ell}^*$ and yet $\pi_n \geq 0$ for all n. Letting $t_n = 1 - \sum_{i=1}^n \ell_i$, the upstream boundary of the n-th firm, we can write the last inequality as

$$p(t_{n-1}) \geq c(\ell_n) + \delta p(t_n), \quad n = 1, \ldots, J.$$

From here we get $p(t_0) \geq c(\ell_1) + \delta\{c(\ell_2) + \delta p(t_2)\}$, and, continuing in this way,

$$p(t_0) \geq \sum_{n=1}^{J} \delta^{n-1} c(\ell_n) + \delta^J p(t_J).$$

In view of this inequality, it will be enough to show that

$$\sum_{n=1}^{J} \delta^{n-1} c(\ell_n) > p^*(1), \tag{12}$$

because then $p(1) = p(t_0) > p^*(1)$, and we have found a contradiction. To establish (12), it is enough to show that $p^*(1)$ is the minimum attained in

$$\min \sum_{n=1}^{\infty} \delta^{n-1} c(\ell_n) \text{ subject to } \{\ell_n\} \subset \mathbb{R}_+ \text{ and } \sum_{n=1}^{\infty} \ell_n = 1, \tag{13}$$

and that the minimizer is unique. In view of proposition 3.3, this is equivalent to stating that the equilibrium allocation $\{\ell_n^*\}$ is the unique solution to (13). Uniqueness is immediate from strict convexity of the objective function, so in fact we need only show that $\{\ell_n^*\}$ solves (13).

For this problem the Karush-Kuhn-Tucker (KKT) conditions for optimality of a feasible allocation $\{\ell_n\}$ are sufficient as well as necessary. The conditions are existence of Lagrange multipliers $\lambda \in \mathbb{R}$ and $\{\mu_n\} \subset \mathbb{R}$ such that

$$\delta^{n-1} c'(\ell_n) = \mu_n + \lambda, \qquad \mu_n \geq 0 \quad \text{and} \quad \mu_n \ell_n = 0 \quad \text{for all} \quad n \in \mathbb{N} \tag{14}$$

Let $\{\ell_n^*\}$ be the Coasian allocation, let $\lambda^* := c'(\ell_1^*)$, and let $\mu_n^* := 0$ for $n = 1, \ldots, N^*$ and $\mu_n^* := \delta^{n-1} c'(0) - \lambda^*$ for $n > N^*$. We claim that $(\{\ell_n^*\}, \lambda^*, \{\mu_n^*\})$ satisfies the KKT conditions. To see this, observe that by repeatedly applying (8) we obtain

$$\delta^{N^*-1} c'(\ell_{N^*}^*) = \delta^{N^*-2} c'(\ell_{N^*-1}^*) = \cdots = \delta\, c'(\ell_2^*) = c'(\ell_1^*) = \lambda^* \tag{15}$$

Now take any $n \in \{1, \ldots, N^*\}$. Since $\mu_n^* = 0$, the first equality in (14) follows from (15) and the second is immediate. On the other hand, if $n > N^*$, then $\ell_n^* = 0$, and hence $\delta^{n-1}c'(\ell_n^*) = \delta^{n-1}c'(0) = \mu_n^* + \lambda^*$, where the last equality is by definition. Moreover, $\mu_n^*\ell_n^* = 0$ as required. Thus is remains only to check that $\mu_n^* = \delta^{n-1}c'(0) - c'(\ell_1^*) \geq 0$. when $n > N^*$.

Since $n > N^*$, it suffices to show that $\delta^{N^*}c'(0) \geq c'(\ell_1^*)$ In view of (15), this claim is equivalent to $\delta^{N^*}c'(0) \geq \delta^{N^*-1}c'(\ell_{N^*}^*)$, or $\delta c'(0) \geq c'(\ell_{N^*}^*)$. Regarding this inequality, recall the definition of $\bar{s}$ as the largest point in $(0,1]$ satisfying $c'(\bar{s}) \leq \delta c'(0)$. In view of lemma 6.3, we have $\ell_{N^*}^* \leq \bar{s}$. Since c' is increasing, we conclude that $\delta c'(0) \geq c'(\ell_{N^*}^*)$ is valid. □

Proof of proposition 3.4. Let $\delta_a \leq \delta_b$. Let T_a and T_b be the corresponding operators. We begin with the claim that $p_a^* \leq p_b^*$. It is easy to verify that if $p \in \mathscr{P}$, then $T_a p \leq T_b p$ pointwise on $[0,1]$. Since T_a and T_b are order preserving (i.e., $p \leq q$ implies $Tp \leq Tq$), this leads to $T_a^n p \leq T_b^n p$. For n sufficiently large, this states that $p_a^* \leq p_b^*$.

Next we show that the number of tasks carried out by the most upstream firm decreases when δ increases from δ_a to δ_b. Let ℓ_i^a be the number of task carried out by firm i when $\delta = \delta_a$, and let ℓ_i^b be defined analogously. Let $N = N_a^*$. Seeking a contradiction, suppose that $\ell_N^b > \ell_N^a$. In that case, convexity of c and proposition 3.2 imply that

$$c'(\ell_{N-1}^b) = \delta_b c'(\ell_N^b) > \delta_a c'(\ell_N^a) = c'(\ell_{N-1}^a).$$

Hence $\ell_{N-1}^b > \ell_{N-1}^a$. Continuing in this way, we obtain $\ell_i^b > \ell_i^a$ for $i = 1, \ldots, N$. But then $\sum_{i=1}^N \ell_i^b > \sum_{i=1}^N \ell_i^a = 1$. Contradiction.

Now we can turn to the claim that $N_b^* \leq N_a^*$. As before, let $N = N_a^*$, the equilibrium number of firms when $\delta = \delta_a$. If $\ell_N^b = 0$, then the number of firms at δ_b is less than $N = N_a^*$ and we are done. Suppose instead that $\ell_N^b > 0$. In view of lemma 6.3, we have $\delta_a c'(0) \geq c'(\ell_N^a)$. Moreover, we have just shown that $\ell_N^a \geq \ell_N^b$. Combining these two inequalities and using $\delta_b > \delta_a$, we have $\delta_b c'(0) \geq c'(\ell_N^b)$. Applying lemma 6.3 again, we see that the N-th firm completes the good, and hence $N_b^* = N_a^*$. □

Lemma 6.8. *If $c(\ell) = e^{\theta\ell} - 1$, then the equilibrium number of firms is given by* (9).

Proof. Let $N = N^*$ be the equilibrium number of firms and let $r := \ln(\delta)/\theta$. From $\delta c'(\ell_{n+1}) = c'(\ell_n)$ we obtain $\ell_{n+1} = \ell_n - r$, and hence $\ell_1 = \ell_n + (n-1)r$. It is easy to check that when $c(\ell) = e^{\theta\ell} - 1$, the constant $\bar{s}$ defined above is equal to r. Applying lemma 6.3 we get $0 < \ell_N \leq r$. Therefore $(N-1)r < \ell_1 \leq Nr$. From $\sum_{n=1}^{N} \ell_n = 1$ and $\ell_1 = \ell_n + (n-1)r$ it can be shown that $N\ell_1 - N(N-1)r/2 = 1$. Some straightforward algebra now yields

$$\frac{1}{2}\left(-1 + \sqrt{1 + 8/r}\right) < N \leq \frac{1}{2}\left(1 + \sqrt{1 + 8/r}\right).$$

The expression for $N = N^*$ in (9) follows. □

References

[1] Kenneth J Arrow. *The organization of economic activity: issues pertinent to the choice of market versus nonmarket allocation*, volume 1, pages 47–64. US Government Printing Office Washington, DC, 1969.

[2] Gary S Becker and Kevin M Murphy. The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics*, 107(4):1137–1160, 1992.

[3] Lawrence M Benveniste and Jose A Scheinkman. On the differentiability of the value function in dynamic models of economics. *Econometrica*, 47(3):727–732, 1979.

[4] Lawrence E Blume, David Easley, Jon Kleinberg, and Eva Tardos. Trading networks with price-setting agents. *Games and Economic Behavior*, 67(1):36–50, 2009.

[5] Alfred D Chandler. *The visible hand: the managerial revolution in American business*. Harvard University Press, 1977.

[6] In-Koo Cho. Competitive equilibrium in a radial network. *RAND Journal of Economics*, pages 438–460, 2003.

[7] Ronald H Coase. The nature of the firm. *Economica*, 4(16):386–405, 1937.

[8] Thibault Fally. On the fragmentation of production in the US. University of Colorado, mimeo, 2011.

[9] John Geanakoplos and Paul Milgrom. A theory of hierarchies based on limited managerial attention. *Journal of the Japanese and International Economies*, 5(3):205–225, 1991.

[10] Christian Ghiglino. Random walk to innovation: why productivity follows a power law. *Journal of Economic Theory*, 147(2):713–737, 2012.

[11] Frank H Hahn. Equilibrium with transaction costs. *Econometrica*, 39(3):417–439, 1971.

[12] Friedrich A Hayek. The use of knowledge in society. *The American Economic Review*, 35(4):519–530, 1945.

[13] Walter P Heller and Karl Shell. On optimal taxation with costly administration. *The American Economic Review*, 64(2):338–345, 1974.

[14] Hugo A Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):1127–1150, 1992.

[15] Boyan Jovanovic. Firm formation with heterogeneous management and labor skills. *Small Business Economics*, 6(3):185–191, 1994.

[16] Takashi Kamihigashi. Elementary results on solutions to the bellman equation of dynamic programming: existence, uniqueness, and convergence. *Economic Theory*, 2013.

[17] Robert E Lucas. On the size distribution of business firms. *The Bell Journal of Economics*, 9(2):508–523, 1978.

[18] Filipe V Martins-da Rocha and Yiannis Vailakis. Existence and uniqueness of a fixed point for local contractions. *Econometrica*, 78(3):1127–1141, 2010.

[19] Filipe V Martins-da Rocha and Yiannis Vailakis. Financial markets with endogenous transaction costs. *Economic Theory*, 45(1-2):65–97, 2010.

[20] Andrew McLennan. Advanced fixed point theory for economics. University of Queensland, mimeo, 2013.

[21] Kieron J Meagher. Generalizing incentives and loss of control in an optimal hierarchy: the role of information technology. *Economics Letters*, 78(2):273–280, 2003.

[22] Douglass C North. Institutions, transaction costs and productivity in the long run. *Washington University in St. Louis, mimeo*, 1993.

[23] Kevin L Reffett. Production-based asset pricing in monetary economies with transactions costs. *Economica*, 63(251):427–443, 1996.

[24] Juan Pablo Rincón-Zapatero and Carlos Rodríguez-Palmero. Existence and uniqueness of solutions to the bellman equation in the unbounded case. *Econometrica*, 71(5):1519–1555, 2003.

[25] Dennis H Robertson. *Control of industry*. Nisbet and Co. Ltd, London, 1923.

[26] Justin Sandefur. On the evolution of the firm size distribution in an African economy. Technical report, CSAE, University of Oxford, 2010.

[27] Patrick W Schmitz. Information gathering, transaction costs, and the property rights approach. *American Economic Review*, 96(1), 2006.

[28] Paul Seabright. *The company of strangers: a natural history of economic life*. Princeton University Press, 2010.

[29] Adam Smith. *An inquiry into the nature and causes of the wealth of nations*. A. and C. Black, 1776.

[30] Oliver E Williamson. Transaction-cost economics: the governance of contractual relations. *Journal of Law and Economics*, 22(2):233–261, 1979.

[31] Oliver E Williamson. The economics of organization: the transaction cost approach. *American Journal of Sociology*, pages 548–577, 1981.

[32] Xiaokai Yang and Yew-Kwang Ng. *Specialization and economic organization: a new classical microeconomic framework*. North-Holland, Amsterdam, 1993.