Federal Reserve Bank of Chicago

Business Networks, Production Chains, and Productivity: A Theory of Input-Output Architecture

Ezra Oberfield

WP 2011-12

Business Networks, Production Chains, and Productivity: A Theory of Input-Output Architecture*

Ezra Oberfield

Federal Reserve Bank of Chicago

ezraoberfield@gmail.com

November 27, 2011

Abstract

This paper studies an analytically tractable model of the formation and evolution of chains of production. Over time, entrepreneurs accumulate techniques to produce their good using goods produced by other entrepreneurs and labor as inputs. The value of a technique depends on both the productivity embodied in the technique and the cost of the particular input; when producing, each entrepreneur selects the technique that delivers the best combination. The collection of known production techniques form a dynamic network of potential chains of production: the input-output architecture of the economy. Aggregate productivity depends on whether the lower cost firms are the important suppliers of inputs. When the share of intermediate goods in production is high, the lower cost firms are selected as suppliers more frequently. This raises aggregate productivity and also increases the concentration of sales of intermediate goods.

Keywords: Networks, Productivity, Supply Chains, Ideas, Diffusion

JEL Codes: O31, O33, O47

*Preliminary and Incomplete, please do not quote without permission of the author. I appreciate the comments of Amanda Agan, Fernando Alvarez, Gadi Barlevy, Marco Bassetto, Jarda Borovicka, Jeff Campbell, Thomas Chaney, Aspen Gorry, Joe Kaboski, Sam Kortum, Alejandro Justiniano, Robert Lucas, Devesh Raval, Rob Shimer, Nancy Stokey, Nico Trachter, and Andy Zuppann, as well as various seminar participants. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System. All mistakes are my own.

Why do some firms have a lower cost of production? One possibility is that a firm is technically efficient: it is good at transforming inputs into output. Alternatively, a firm could have access to a supplier that provides particularly cheap inputs. When inputs are cheap, output is inexpensive to produce even if production is not efficient. Having a low cost of production depends not just on what the firm knows, but on who it knows. If finding a low-cost supplier is important, much of what we think of as aggregate productivity is comprised of relationships that link entrepreneurs.

This paper develops a model of the formation and evolution of input-output links. In the model, entrepreneurs search for the most cost-effective techniques to produce their goods and for new uses for their goods. It is based on the premise that there may be multiple ways to produce a good, each using a different set of inputs.[1] When one entrepreneur meets another, she may develop a new technique for producing her good using the other entrepreneurs's good as an input. When deciding which technique to use, an entrepreneur cares both about the technical efficiency of the technique and about the cost of the inputs. If she can charge a lower price, other entrepreneurs are more likely to use her good as an input. As new techniques are discovered and as suppliers adjust their prices, entrepreneurs substitute across methods of production.

The collection of known production techniques form a dynamic network comprising the set of supply chains available to make each final product. When an entrepreneur discovers a new, more cost effective technique, the cost savings diffuse through the network. How large those effects are and how quickly these more efficient techniques are developed hinge on both the density and structure of the network.[2]

[1] For example, an electricity producer may have a production technique that uses oil as an input and another less productive technique that uses coal.

[2] Consider an entrepreneur who discovers an extremely cost effective technique. If the firm has many potential customers (i.e., there are many other firms that have techniques that use the entrepreneur's good as an input), those lower prices will ripple down many supply chains. If, however, the firm has few potential customers, the lower prices may not have much of an immediate impact.

I first study a planner's problem in an economy in which new techniques arrive randomly to all entrepreneurs at a uniform rate. For this baseline case, there is a closed form expression for aggregate productivity, relating features of the network (and parameters that shape the network) to aggregate output. An important determinant of aggregate productivity is whether the lower cost producers are able to become the important suppliers of the economy. To become an important supplier, a firm must have many potential customers who know how to use the firm's good and to be able to charge those potential customers a relatively low price. When intermediate goods are more important in production (relative to labor), the ability to charge a low price becomes more important in winning customers, and the lowest cost firms are more likely to become superstar suppliers. This increases aggregate productivity and also increases the market concentration in sales of intermediate goods.

To better understand what determines the emergence of superstar suppliers, I study a simple extension that allows for heterogeneous rates of link formation across different subgroups. The particular configuration determines how productivity and demand covary and hence how frequently the lowest cost firms are selected to supply intermediate goods. Concentration of link formation among a subset of firms, e.g., in a city, leads to positive correlation of productivity and demand across subgroups and hence higher aggregate output.

A final extension allows for production functions that use multiple inputs. Raising the number of inputs makes the cross sectional distribution of marginal cost more concentrated and dampens the impact of superstar suppliers.

I then study equilibria of a particular market structure with monopolistic competition in sales of final goods and bilateral two part pricing between all potential input-output pairs. I consider the set of pairwise stable equilibria: arrangements for which no there are no profitable unilateral or pairwise deviations. For each of these equilibria, all firms set the marginal price equal to marginal cost (along a non-negative fixed fee) so production chains avoid double marginalization. There are many equilibria that decentralize the planner's optimal allocation but differ in how profits are divided across firms. There is also always at least one equilibrium in which no production occurs (although there is a sense in which that equilibrium is fragile).

This paper relates to several disparate literatures. Most directly connected is the literature on social networks, especially as applied to firms. Atalay et al. (2011) document cross-sectional facts about input-output links, while Chaney (2011) discusses the network structure of international shipments of goods. Both explain their empirical findings using extensions of the preferential attachment models of Barabasi and Albert (1999) and Jackson and Rogers (2007). Relative to this literature, this paper makes three contributions. First, this paper endogenizes the formation of observed input-output links in the economy. Each entrepreneur selects a supplier from among many potential suppliers, and these choices are mediated by prices. In the preferential attachment model, the formation of links is mechanical: no choices are made. The endogeneity of network formation is advantageous in that it opens the door to the analysis of policy experiments that change incentives and thus the link formation process. Second, in this model the network structure determines not just the distribution of input-output links but also both firm level and aggregate productivity, providing a more clear connection between the network structure and economic variables of interest such as aggregate output and welfare. Third, this paper gives a natural alternative explanation for a skewed cross-sectional distribution of input-output links.

The model emphasizes technological interdependence in the spirit of Rosenberg (1979)[3]. In

[3] "The social payoff of an innovation can rarely be identified in isolation. The growing productivity of industrial economies is the complex outcome of large numbers of interlocking, mutually reinforcing technologies, the individual components of which are of very limited economic consequence by themselves. The smallest relevant unit of observation is seldom a single innovation but, more typically, an interrelated clustering of innovations."

trying to make sense of the black box that is aggregate productivity, this paper provides a structure that connects features of technological relationships to an aggregate production function and to cross sectional characteristics of firms.

The structure of the model is also related to the work of Kortum (1997), Eaton and Kortum (2002), Alvarez et al. (2008), and Lucas (2009) who study flows of ideas. The most closely connected is Lucas (2009), who studies a model in which individuals learn ideas from other individuals over time.

The idea that network structure determines how shocks propagate through an economy has been incorporated into the real business cycle literature, centered around the model of Long and Plosser (1983). Recently, the discussion has focused on whether shocks to particularly well connected sectors can account for aggregate fluctuations.[4] These models typically assume that each sector has a representative firm that produces using a Cobb-Douglas production function, using inputs from all other sectors.[5] Like these models, this paper is concerned with how cost savings spill over to other firms through input-output links. However, while that literature has focused on the consequences of sectoral shocks taking the input-output structure as given, this paper studies the endogenous formation and evolution of the input-output structure.

This paper proceeds as follows: Section 1 describes the basic technology, setting up and solving a social planner's problem. In this section there are simple formulas relating the density of the network to aggregate output. Section 2 describes the size distribution. Section 3 discusses market structures that could be overlaid on the economic environment. Section 4 generalizes the model

[4] See Horvath (1998), Dupor (1999), Carvalho (2007), Acemoglu et al. (2010), and Foerster et al. (2008). This literature has focused on the sectoral level both because the most fine input-output data is at that level and because solving these models involves inverting matrices, which becomes computationally intensive as the number of nodes in the network grows large.

[5] Jones (2008) uses a similar model to argue the input-output structure can help us understand cross country income differences. In that setup misallocation in one sector raises input prices in other sectors, and the magnitude of the overall effect depends on the input-output structure.

presented in the first section to multiple types to allow for more interesting network configurations, while Section 5 studies how particular configurations relate to aggregate output. Section 6 extends the model to allow for multiple inputs. Section 7 concludes.

1 The Baseline Model

1.1 Economic Environment

There is a unit mass of infinitely-lived firms, and each firm is associated with producing a particular good. Each good is used for final consumption and potentially can be used as an intermediate input by other firms. A representative consumer has Dixit-Stiglitz preferences over the goods and supplies labor inelastically (both of these can easily be relaxed). There is no storage technology.

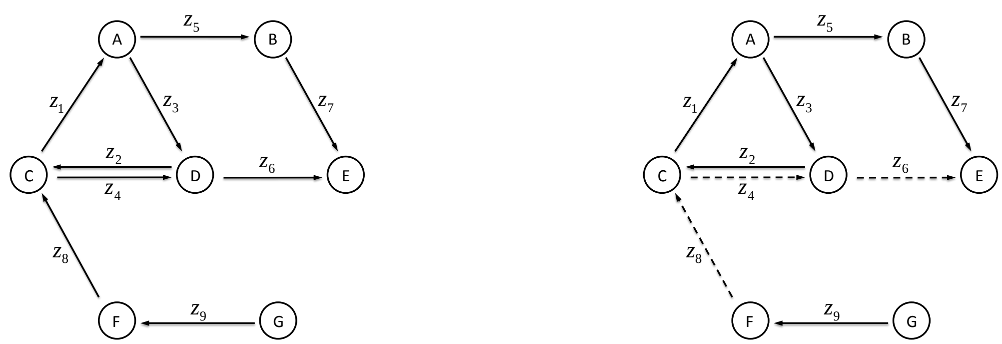
Over time, each firm accumulates production techniques. For a firm, a technique is method of producing its good using some other firm's good as an intermediate input. Each technique $\phi = \{j, i, z\}$ consists of three components: (i) the good that is produced, j; (ii) the good used as an input, i; and (iii) a production technology associated with using that input, indexed by the productivity parameter z:

$$y = \frac{1}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}} z x^{\alpha} L^{1-\alpha}$$

where y is the quantity of output of good j produced, x is the quantity of good i used as an input, L is labor.

Time is continuous and firms discover and lose access to techniques randomly over time. Each firm discovers new techniques at the arrival rate $\tilde{\mu}(t)$. When firm j discovers a new technique, $\phi = \{j, i, z\}$ the identity of the supplier i is random and uniformly distributed across all firms in the economy.[6] The productivity parameter, z, of a technique is drawn from a fixed distribution with CDF $H(\cdot)$ and is constant over time. It is assumed that the support of H is bounded below by some $z_0 > 0$ and that $\int_{z_0}^{\infty} \log(z) dH(z) < \infty$. Each existing technique becomes infeasible at rate δ.[7]

Given a menu of techniques and input prices, each firm produces using the technique that delivers the best combination of input cost and productivity. Since the production functions exhibit constant returns to scale, generically using a single technique is optimal.[8]



(a) The Set of Techniques

(b) Techniques that are Used

Figure 1: A Graphical Representation of the Input-Output Structure

Figure 1a shows the set of techniques at a point in time. Each node is a firm, and an edge that connect two nodes corresponds to a technique. The direction indicates which firm produces the output, and which firm produces the output. Each edge has a weight associated with the productivity of the technique. In Figure 1b the solid arrows are techniques that are currently used, while the dashed arrows correspond to techniques that are not.

At a point in time, the state of the economy can be summarized by the set of available techniques, $\Phi(t)$. Figure 1 gives a visual representation of this set as a weighted, directed graph. In the figure,

[6] Section 4 relaxes this uniformity.

[7] δ plays a minor role in the analysis and setting $\delta = 0$ would change little. It is included (i) for generality and (ii) so that when $\tilde{\mu}(t)$ is constant there is a well defined steady state.

[8] In principle a firm could produce using more than one technique. In the analysis I will assume without loss that each firm uses only a single technique.

firms are represented by nodes. Each technique is represented by an edge or link connecting two nodes. Each edge has a direction that corresponds to the flow of goods for the technique, indicating which firm provides the intermediate input and which produces the output. In addition, each edge has a weight (a number) corresponding to the productivity of the technique.

Figure 1a shows the set of all available techniques. Note that several firms, C, D, and E, each have knowledge of multiple techniques that use different producers. In equilibrium they will each select a single technique with which to produce. Figure 1b gives an example of these selections.

Over time the network will evolve as new techniques are discovered and some old techniques become infeasible. While the productivity of a technique is constant over time, the attractiveness of a technique varies with the cost of the associated inputs.

1.2 A Planner's Problem

Throughout Section 1 I will focus on a planner's problem in order to build intuition about the economic environment and to describe solution techniques without getting bogged down with the details of a particular market structure. There are many market structures that could be layered on top of the technological environment. In Section 3 I discuss a particular market structure that decentralizes the planner's solution.

Consider the problem of a planner that takes the network of techniques as given. At each point in time the planner makes production decisions and allocates labor to maximize the instantaneous utility of the representative agent. To increase readability time subscripts will be suppressed when unnecessary. Let y_j^0 be production of good j for final consumption. Given the set of existing techniques, Φ, we can define two subsets relevant for each firm: let Φ_j^S be the set of techniques available to produce good j (potential suppliers for firm j) and let Φ_j^B be the set of techniques that would use good j as an input (potential buyers of good j). For a technique $\phi = \{j, i, z\}$, define the following quantities:

- $z(\phi)$ is the productivity parameter associated with the technique.
- $s(\phi)$ is the identity of the supplier, in this case i.
- $b(\phi)$ is the identity of the buyer, in this case j.

The planner selects a technique for firm j to use, ϕ_j^*, from among the available techniques to produce good j, Φ_j^S. With that, the planner chooses how much of good j to produce, y_j, a quantity of labor, L_j, and the quantity of inputs of good $s(\phi_j *)$ to use.

Formally, the planner chooses an allocation $\left\{\phi_j^*, y_j^0, y_j, x_j, L_j\right\}_{j\in J}$ to maximize final consumption:

$$\max \left(\int_{j\in J} (y_j^0)^{\frac{\varepsilon}{\varepsilon-1}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$$

subject to: (i) technological constraints

$$y_j \leq \frac{1}{\alpha^\alpha(1-\alpha)^{1-\alpha}} z(\phi_j^*) x_j^\alpha L_j^{1-\alpha}, \qquad \forall j \in J$$

(ii) goods feasibility constraints

$$y_j^0 + \sum_{\phi\in\Phi_j^B|\phi=\phi_{b(\phi)}^*} x_{b(\phi)} \leq y_j \qquad \forall j \in J$$

and (iii) a labor resource constraint

$$\int_{j\in J} L_j \leq L$$

The left hand side of the second constraint for good j consists of the uses of the good: output for final consumption and for use as an intermediate input in other firms' production.

Let MC_j be the marginal social cost of producing good j (the multiplier on the goods feasibility constraint for j), and let w be the marginal social cost of labor (the multiplier on the labor resource constraint). The first order necessary conditions from this problem imply that for each j,

$$\frac{MC_j}{w} = \min_{\phi \in \Phi_j^S} \frac{1}{z(\phi)} \left(\frac{MC_{s(\phi)}}{w} \right)^{\alpha} \tag{1}$$

For each technique ϕ, $\frac{1}{z(\phi)} \left(\frac{MC_{s(\phi)}}{w} \right)^{\alpha}$ gives the marginal social cost of producing good j using that technique, with the planner using the one that delivers the lowest marginal social cost.

It will be convenient to define $q_j \equiv \frac{1}{MC_j}$ as a measure of the efficiency of producing good j. If we choose units of utility so that $w = 1$, we can rewrite equation (1) as

$$q_j = \max_{\phi \in \Phi_j^S} z(\phi) q_{s(\phi)}^{\alpha} \tag{2}$$

1.3 The Supply Chain Interpretation

Given the structure of the network, we can back out the production technology used by the social planner to produce each good. We consider here how labor is allocated across the different stages of production in the supply chain for each good.

For a given good, say j, the planner uses a particular chain of techniques to produce good j. Let $\{\phi_0 \phi_1 \phi_2 ...\}$ denote the chain of techniques for good j, with ϕ_0 being furthest downstream (the technique actually used by firm j).

Let $\bar{L}_j$ be total amount of labor used in production of good j for final consumption (given constant returns to scale in production, this is well defined). Also let L_j^k be the labor used in the kth to last stage of production of good j for final consumption, so that L_j^0 is the labor used by firm j to produce for final consumption (as opposed to for production of the good for intermediate use). Aggregating across stages gives $\bar{L}_j = \sum_{k=0}^{\infty} L_j^k$.

The first order conditions imply that $L_j^{k+1} = \alpha L_j^k$: the labor used at each stage is a constant fraction of the labor used in the subsequent stage. We can therefore write

$$\bar{L}_j = \sum_{k=0}^{\infty} L_j^k = \sum_{k=0}^{\infty} \alpha^k L_j^0 = \frac{1}{1-\alpha} L_j^0$$

We also have the first order condition $y_j^0 = \frac{1}{1-\alpha} q_j L_j^0$, so that

$$y_j^0 = q_j \bar{L}_j \tag{3}$$

In a sense equation (3) should not be surprising; q_j was defined to be the ratio of the marginal social cost of producing good j and the marginal social cost of labor, or the efficiency with which the planner can produce good j in units of labor.

We can derive a more basic interpretation of equation (3) in terms of more fundamental objects in the model. With similar notation, we can define q_j^k to be the efficiency of the kth to last firm in the chain of production for good j, and $z(\phi_k)$ to be the productivity parameter of the technique in the kth step. This means that along the supply chain for good j, $q_j^k = z(\phi_k)\left(q_j^{k+1}\right)^{\alpha}$. By definition, $q_j^0 = q_j$, so we make repeated substitutions to get

$$\begin{aligned} q_j &= q_j^0 = z(\phi_0)\left(q_j^1\right)^{\alpha} = z(\phi_0)\left[z(\phi_1)\left(q_j^2\right)^{\alpha}\right]^{\alpha} = \ldots \\ &= \prod_{k=0}^{\infty} z(\phi_k)^{\alpha^k} \end{aligned}$$

The planner therefore faces a production function describing the social cost of producing good final j:

$$y_j^0 = \left[\prod_{k=0}^{\infty} z(\phi_k)^{\alpha^k}\right] \bar{L}_j \tag{4}$$

The efficiency with which the planner can produce good j depends on the productivity of each technique at each step in the supply chain, with the techniques furthest downstream weighted more heavily.

1.4 The Allocation of Labor and Welfare

We now use several more first order conditions from the planners problem to arrive at an expression for total final consumption. Define $Q \equiv \left(\int_J q_j^{\varepsilon-1}\right)^{\frac{1}{\varepsilon-1}}$, a standard productivity aggregator for economies with Dixit-Stiglitz preferences. The first order conditions with respect to each y_j^0 imply $\frac{y_j^0}{Y^0} = \left(\frac{q_j}{Q}\right)^{\varepsilon}$. We can now use the labor resource constraint and equation (3) to write

$$L = \int_J \bar{L}_j = \int_J \frac{y_j^0}{q_j} = \int_J Y^0 Q^{-\varepsilon} q_j^{\varepsilon-1} = Y^0/Q \tag{5}$$

or more conveniently

$$Y^0 = QL \tag{6}$$

To solve for aggregate output, we need to characterize firms' efficiencies. One could use equation (2) to create an operator on $\{q_j\}_{j \in J}$ and look for a fixed point.[9] However, with a continuum of firms, this is neither computationally feasible nor would it be particularly illuminating. We proceed to impose more structure on the set of techniques motivated by the dynamics of the model.

1.5 The Cross Sectional Distribution of Efficiency

To this point, the treatment of the network of techniques has been quite general. In this section we impose the probabilistic structure of the dynamic model in order to more clearly characterize the solution to the planner's problem. In particular, we will set up a fixed point problem for the cross sectional distribution of efficiency.

There are two important elements for characterizing the distribution. The first is how techniques are distributed across firms. The second is the distribution of efficiency delivered by each technique. We characterize each in turn.

1.5.1 The Distribution of Techniques

Here we describe how techniques are distributed across firms and how this distribution evolves. Roughly, given the history of Poisson arrival rates of new techniques and decay rates of existing techniques, the fraction of firms with n techniques at time t can be fully described by a Poisson distribution with a mean $\tilde{\lambda}(t)$.

More formally, let $\omega(n, t)$ be the fraction of firms with access to n techniques at time t. Over

[9] Taking logs of both sides gives an operator T, where the jth element of $T(\{\log q_j\}_{j \in J})$ is $\max_{\phi \in \Phi_j^B} \{\log z(\phi) + \alpha \log q_{s(\phi)}\}$. T satisfies monotonicity and discounting. If the support of z were bounded above, the operator would be a contraction on the appropriate bounded function space. If the efficiencies $\{q_j\}_{j \in J}$ can take values on the extended real line, then the mapping will not be a contraction. Indeed, there may be multiple fixed points, as discussed below.

time, this evolves according to the following law of motion:

$$\dot{\omega}(n,t) = \tilde{\mu}(t)\omega(n-1,t) + (n+1)\delta\omega(n+1,t) - \tilde{\mu}(t)\omega(n,t) - n\delta\omega(n,t) \tag{7}$$

$\omega(n,t)$ increases when a firm with $n-1$ techniques discovers a new one and when a firm with $n+1$ techniques loses one of their $n+1$ techniques. Similarly $\omega(n,t)$ decrease when a firm with n techniques either gains a new technique or loses one of its n techniques.

If at some t_0 the distribution of n is given by a Poisson distribution with mean $\tilde{\lambda}(t_0)$, then a solution to equation (7) is such that at any time $t > t_0$, the distribution of n will also follow a Poisson distribution with mean $\tilde{\lambda}(t)$, where $\tilde{\lambda}(t)$ satisfies the differential equation:[10]

$$\dot{\tilde{\lambda}}(t) = \tilde{\mu}(t) - \delta\tilde{\lambda}(t) \tag{8}$$

To interpret this, it helps to take the limit as $t_0 \to -\infty$ (and imposing that $\lim_{t_0 \to -\infty} \tilde{\lambda}(t_0)$ is bounded) giving

$$\tilde{\lambda}(t) = \int_{-\infty}^{t} e^{-(t-\tau)\delta}\tilde{\mu}(\tau)d\tau$$

This is closely related to the fact that the sum of independent Poisson random variables is also a Poisson random variable. The takeaway from this is that regardless of history of arrival rates $(\{\tilde{\mu}(\tau)\}_{\tau \leq t})$, the cross-sectional distribution of techniques $\omega(n,t)$ at a given point in time can be summarized by a single number, $\tilde{\lambda}(t)$. Since $\tilde{\lambda}(t)$ is also the average number of techniques per firm, I will refer to it as the *density* of techniques in the network.

[10] In fact, for an arbitrary initial distribution, the distribution of links will converge asymptotically to Poisson distribution.

1.5.2 The Cross Sectional Distribution of Efficiency

Let $F(q)$ be the fraction of firms with efficiency no greater than q given the decisions of the planner. This is an endogenous object that will need to be solved for. The strategy is to use the fact that each potential supplier has efficiency no greater than q with probability $F(q)$ to set up a fixed point problem for F.

If firm j discovers a single technique ϕ, there are two parts that determine how useful it is: a productivity parameter, z, drawn from an exogenous distribution $H(z)$, and the efficiency of the supplier, q_i. Recall from equation (2) that if firm j produces using technique $\phi = \{j, i, z\}$, j will produce at efficiency $q_j = z(\phi)q_i^\alpha$. Let $G(q)$ be the cumulative distribution of the efficiency delivered by a single random technique. Given equation (2), we can write $G(q)$ as

$$G(q) = \int_0^\infty F\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right) dH(z) \tag{9}$$

To interpret this, note that for each z, $F\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right)$ is the portion of potential suppliers that, in combination with that z, would leave the firm with efficiency no greater than q.

We now ask, what is the probability that, given all of its techniques, a firm has efficiency no greater than than q? We can write this as

$$\begin{aligned}
\Pr\left(q_j \leq q\right) &= \sum_{n=0}^{\infty} \Pr\left(\text{All } n \text{ draws are} \leq q\right) \omega\left(n\right) \\
&= \sum_{n=0}^{\infty} G\left(q\right)^n \frac{\tilde{\lambda}^n e^{-\tilde{\lambda}}}{n!} \\
&= e^{-\tilde{\lambda}[1-G(q)]}
\end{aligned}$$

To interpret this last expression, if $\tilde{\lambda}\,[1-G(q)]$ is a parameter of a Poisson distribution (the arrival rate of techniques that would provide efficiency better than q), then $e^{-\tilde{\lambda}[1-G(q)]}$ is the probability that no such techniques arrived.

When the number of firms is large, a standard abuse of the law of large numbers gives $\Pr(q_j \le q) = F(q)$.[11] We can substitute the expression for $G(q)$ from equation (9) to get a fixed point problem for the distribution of efficiency $F(q)$:

$$F(q) = e^{-\tilde{\lambda}\int_0^\infty \left[1-F\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right)\right] dH(z)} \tag{10}$$

This recursive equation is the key to solving the planner's problem.

1.6 Properties of the Planner's Solution

Consider the space $\bar{\mathcal{F}}$ of non-decreasing functions $f : \mathbb{R}^+ \mapsto [0,1]$. Consider the operator T on this space defined as

$$Tf(q) \equiv e^{-\tilde{\lambda}\int_0^\infty \left[1-f\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right)\right] dH(z)}$$

With probability one the planner's solution will be a fixed point of the operator T on $\bar{\mathcal{F}}$.

The qualitative behavior of network depends on whether the average number of techniques $\tilde{\lambda}$ is greater or less than 1. The more interesting case in which the average number of links exceeds one will be the focus of this paper, but for completeness I will discuss both.

[11] The proof of Proposition 1 uses such a law of large numbers for a continuum of random variables described by Uhlig (1996). To use this one must verify that firms' efficiencies are pairwise uncorrelated. In the present context this is not immediately obvious: it is possible that two firms' supply chains overlap or that one is in the other's supply chain. However, by assumption the network is sufficiently sparse that with high probability the supply chains will not overlap: there is a continuum of firms but only a countable number of those are in any of given firm's potential supply chains. Therefore, for any two firms, the probability that their supply chains overlap is zero. The law of large numbers then implies that equation (10) holds with probability one.

If firm j does not have access to any techniques, it cannot produce. Similarly, if firm j has techniques but its suppliers do not, then those suppliers will not be able to produce and consequently neither will firm j. Continuing with this logic, if a supply chain is finite, it is not viable.

Consider the probability that a single firm will have access to at least one supply chain that continues indefinitely; only firms with access to such a supply chain will have positive efficiency. As shown in Appendix B the probability that a firm has no such chains is the smallest root ρ of $\rho = e^{-\tilde{\lambda}(1-\rho)}$.[12] For $\tilde{\lambda} \leq 1$, the probability of no such chains is one, while for $\tilde{\lambda} > 1$, the probability is strictly less than one.[13]

Few Techniques: $\tilde{\lambda} \leq 1$

If $\tilde{\lambda} \leq 1$, there are so few techniques available that the probability that any individual firm has access to a viable supply chain is zero. In this case one can show that T is a contraction, with the unique solution $f = 1$: all firms produce with efficiency 0.

Many Techniques: $\tilde{\lambda} > 1$

As $\tilde{\lambda}$ crosses the critical value of 1, one can show that there are multiple fixed points of the operator T on $\bar{\mathcal{F}}$ (see Appendix C). Recall that the functional equation equation (10) was constructed from necessary (but not sufficient) conditions to the planner's problem, so one must check which of these solutions to equation (10) actually solves the planner's problem.

There are two solutions in which $F(q)$ is constant for all q, both of which stem from the fact that equation (10) is formulated as a recursive equation. The first is $F(q) = 1$ for all q, which again corresponds to zero efficiency (infinite marginal social cost) for all goods. However, the rationale

[12] There is a direct analogy between this problem and the Galton-Watson problem. Consider a world in which every individual has a random number of children drawn from some fixed distribution. The Galton-Watson problem is: What is the asymptotic probability of extinction?

[13] The starkly different behavior of the network when $\tilde{\lambda}$ crosses 1 is called a phase transition. Such a phase transition is a typical property of random graphs, a result associated with the Erdos-Renyi Theorem. See Kelly (1997) and Kelly (2005) for examples in which this kind of phase transition is given an economic interpretation.

is different than when $\tilde{\lambda} \leq 1$; here, the logic is recursive. If the marginal social cost of every input is infinite, then the marginal social cost of each output must be infinite as well. The allocation that arises from this solution is feasible, but we will show that it is dominated by another feasible allocation and is therefore not the solution to the planner's problem.

There is a second constant solution, $F(q) = \rho \in (0, 1)$, which follows a similar recursive logic. This fixed point implies infinite marginal social cost for those firms that cannot produce, and zero marginal social cost for all other firms. In other words, any firm with a viable supply chain has infinite efficiency. The rationale is similar: if inputs have zero marginal social cost, output has zero social cost. Unfortunately, this leads to an infeasible allocation, and is therefore not a solution to the planners problem either.[14]

There is always a third fixed point. In Appendix C a subset $\mathcal{F} \subset \bar{\mathcal{F}}$ is constructed along with a partial ordering. The subset does not contain either of the two constant fixed points. We can show that there is a fixed point of T using the Tarski fixed point theorem, which also provides an algorithm to numerically solve for such a fixed point. Further we can show that the fixed point is unique. See Appendix C for a more complete statement of the theorem and a proof.

Proposition 1 *There exists a unique fixed point of T on $\mathcal{F}$, F^{sp}. With probability one, F^{sp} is the CDF of the cross sectional distribution of efficiencies in the solution to the planner's problem and aggregate productivity is $Q = \left(\int_0^\infty q^{\varepsilon-1} dF^{sp}(q)\right)^{\frac{1}{\varepsilon-1}}$.*

Multiple solutions to first order necessary conditions of the planner's problem is actually a feature of most models in which a portion of output is used simultaneously as input, such as

[14] These two constant solutions have further economic meaning. Given the distribution of productivity draws $H(\cdot)$ with support $[\underline{z}, \bar{z}]$ with $0 \leq \underline{z} \leq \bar{z} \leq \infty$, let $\underline{q}$ and $\bar{q}$ be the lowest and highest possible efficiencies among firms that are able to produce. $\underline{q}$ ($\bar{q}$) is derived from the supply chain in which every technique has the worst (best) possible productivity draw, so that $\underline{q} = \underline{z}^{\frac{1}{1-\alpha}}$ ($\bar{q} = \bar{z}^{\frac{1}{1-\alpha}}$). If $\tilde{\lambda} > 1$ the solution to the planners problem must have $F(\bar{q}) = 1$ and $F(\underline{q}) = \rho$, the two constant solutions to equation (10).

a standard growth model with roundabout production. If the same good enters a technological constraint as both an input and an output, the Lagrange multiplier on that good will be on both sides of a first order condition. Consequently the first order condition will be satisfied if the Lagrange multiplier takes the value of zero or infinity. One can usually sidestep this issue by finding an alternative way to describe the production technology, e.g., solving for final output as a function of primary inputs. Much of the work in the proof of Proposition 1 is in finding and characterizing such an alternative description of production possibilities.

1.7 A Parametric Assumption

I now describe a special case that proves to be analytically tractable. Assume that the productivity parameter embodied in a technique is drawn from a Pareto distribution, $H(z) = 1 - \left(\frac{z}{z_0}\right)^{-\zeta}$, with the restriction that $\zeta > \varepsilon - 1$ so that final output is finite. In addition, parameterize the arrival rate (and initial condition) of techniques so that $\tilde{\mu}(t) = \mu(t) z_0^{-\zeta}$ for all t which implies that $\tilde{\lambda}(t) = \lambda(t) z_0^{-\zeta}$.

With these assumptions we can compare economies with different values of z_0 (holding $\{\lambda(t)\}$ fixed). In an economy with a lower z_0, there will be two differences: and (ii) techniques have stochastically lower productivity and (ii) each firm discovers new techniques more frequently. In fact, the parameterization is such that varying z_0 has no impact on the average number of techniques with productivity above any threshold $\hat{z}$, $\tilde{\lambda}\left(1 - H(\hat{z})\right)$: the two effects cancel exactly. The only difference is that with a lower z_0 there are additional relatively unproductive techniques.

We then look at the limit of a sequence of economies as $z_0 \to 0$. This adds many relatively unproductive techniques (low z) to the economy without changing the number of productive techniques (high z). In the limit, the measure of firms without access to any techniques goes to zero.

In this special case, we can show that every solution $F(\cdot)$ to equation (10) follows a Frechet distribution. To see this, note that we can use the change of variables $x = (q/z)^{1/\alpha}$ to write

$$
\begin{aligned}
1 - G(q) &= \int_{z_0}^{\infty} H'(z) \left(1 - F\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right)\right) dz \\
&= \int_{0}^{\left(\frac{q}{z_0}\right)^{\frac{1}{\alpha}}} H'\left(\frac{q}{x^{\alpha}}\right) (1 - F(x))\, q\alpha x^{-\alpha-1} dx
\end{aligned}
$$

Using the functional form $H'(z) = \zeta z_0^{\zeta} z^{-\zeta-1}$, we can then write

$$
\begin{aligned}
\tilde{\lambda}\left[1 - G(q)\right] &= \lambda z_0^{-\zeta} \int_{0}^{\left(\frac{q}{z_0}\right)^{\frac{1}{\alpha}}} \zeta z_0^{\zeta} \left(\frac{q}{x^{\alpha}}\right)^{-\zeta-1} (1 - F(x))\, q\alpha x^{-\alpha-1} dx \\
&= q^{-\zeta} \lambda \int_{0}^{\left(\frac{q}{z_0}\right)^{\frac{1}{\alpha}}} \alpha\zeta x^{\alpha\zeta-1} (1 - F(x))\, dx
\end{aligned}
$$

For any $F(\cdot)$, as $z_0 \to 0$, this expression will clearly go to $q^{-\zeta}$ multiplied by a constant. Label this constant θ, so that equation (10) can be written as $F(q) = e^{-\theta q^{-\zeta}}$, the cumulative distribution of a Frechet random variable. Note that the exponent ζ is the same as that of the Pareto distribution H. This means that the distribution of efficiencies F inherits the tail behavior of the distribution of productivity draws, H.

We next solve for θ, which was defined to satisfy

$$
\theta = \lambda \int_{0}^{\infty} \alpha\zeta x^{\alpha\zeta-1} (1 - F(x))\, dx
$$

Integrating by parts gives

$$
\theta = \lambda \int_{0}^{\infty} x^{\alpha\zeta} F'(x) dx
$$

Plugging in the functional form $F(q) = e^{-\theta q^{-\zeta}}$ and making the substitution $s = \theta x^{-\zeta}$ gives

$$\theta = \lambda \int_0^\infty \theta^\alpha s^{-\alpha} e^{-s} ds$$

so that θ satisfies[15]

$$\theta = \Gamma(1-\alpha)\lambda\theta^\alpha \tag{11}$$

where $\Gamma(\cdot)$ is the gamma function.

With this, we can compute Q, the relevant measure of welfare:

$$Q = \left(\int_0^\infty q^{\varepsilon-1} dF(q)\right)^{\frac{1}{\varepsilon-1}} = \theta^{1/\zeta}\Gamma\left(1 - \frac{\varepsilon-1}{\zeta}\right)^{\frac{1}{\varepsilon-1}}$$

Putting these together, we get an expression for final consumption:

$$Y = \Gamma\left(1 - \frac{\varepsilon-1}{\zeta}\right)^{\frac{1}{\varepsilon-1}} \Gamma(1-\alpha)^{\frac{1}{1-\alpha}\frac{1}{\zeta}} \lambda^{\frac{1}{1-\alpha}\frac{1}{\zeta}} L \tag{12}$$

There are several immediate implications. First, aggregate output is increasing in the density of the network, λ. In a more densely populated network, firms on average have a larger set of supply chains to choose from, and hence are more likely to have lower cost.

Second, the share of intermediate goods in production α plays two roles. First, as in other models with roundabout production, it determines the extent to which lower prices of input cost feed back into lower cost of production. This is the exponent $\frac{1}{1-\alpha}$ that appears in several places. Second, α determines the likelihood that the lowest cost producers are selected as the actual suppliers of

[15] I write equation (11) in this form rather than solving directly for θ in order to emphasize the fact that the equation has three non-negative roots, two of which are zero and infinity. In addition, in later sections it will be easier to see parallels with the analogous expressions when this equation when written in this form.

inputs. Recall that the efficiency delivered by a single technique is zq^{α}, where z is the productivity embedded in the technique and q is the efficiency of supplier. α determines the relative impact of each of these factors on the cost effectiveness of a technique, and consequently in the selection of supplier. When α is high, the cost of the inputs matters more, which means the most efficient producers are selected as suppliers more frequently. In other words, superstar suppliers will be much more relevant for aggregate production when α is closer to one. Mathematically, this shows up in the term $\Gamma(1-\alpha)$.[16] Summarizing, when α is high, each supplier is able to pass through cost savings to its customers at a higher rate *and* the most efficient firms are selected to be suppliers more frequently, so that their high efficiency can be passed through to a larger share of customers. This will be discussed in greater detail in Section 2

1.8 Dynamics

Over time, some firms discover new techniques while others substitute across techniques in response to changes in cost along their supply chains. For aggregate quantities, however, all of these changes can be summarized by changes in the density of the network, λ. All relevant aggregate dynamics can be summarized by two equations

$$Y = \kappa \lambda^{\frac{1}{1-\alpha}\frac{1}{\zeta}} L$$

and

$$\dot{\lambda}(t) = \mu(t) - \delta\lambda(t)$$

If the (normalized) arrival rate of new techniques $\mu(t)$ is constant over time (and if $\delta > 0$), then

[16] $\Gamma(x)$ is decreasing on $(0, 1)$. $\Gamma(1) = 1$ and $\lim_{x\to 0} \Gamma(x) = \infty$.

there is a steady state with

$$\lambda_{ss} = \frac{\mu}{\delta}$$

Alternatively if the arrival rate of new techniques is growing over time, say $\mu(t) = \mu e^{\gamma t}$, then there is a balanced growth path (for any $\delta \geq 0$) with

$$\frac{\dot{\lambda}}{\lambda} = \gamma \qquad \text{and} \qquad \frac{\dot{Y}}{Y} = \frac{1}{1-\alpha}\frac{1}{\zeta}\gamma$$

2 Size Distribution

Here I discuss two dimensions of the cross sectional distribution of size. First, I give expressions for the conditional and unconditional distributions of the number of customers (other firms that purchase intermediate goods). Second, I describe the cross sectional distribution of employment.

2.1 Number of Customers

Consider a single draw of a technique that uses firm i's good as an input. Given the firm's efficiency, q_i, we can compute the probability that the technique is the potential buyer's best available technique. To do this, we first characterize the following object: For a potential customer that has drawn a technique that uses i, what is the probability that it has no other techniques better that deliver efficiency better than q?

Given the Poisson distribution over the number of techniques, a firm will have $n-1$ other techniques with probability $\frac{e^{-\tilde{\lambda}}\tilde{\lambda}^n}{n!(1-e^{-\tilde{\lambda}})}$. The CDF of efficiency delivered by each of these techniques is $G(q)$. We can therefore write the probability that the potential buyer has no other technique that delivers better than q, conditioning on having at least one technique, as:

$$\frac{\sum_{n=1}^{\infty}\frac{e^{-\tilde{\lambda}}\tilde{\lambda}^n}{n!}G(q)^{n-1}}{1-e^{-\tilde{\lambda}}} = \frac{1}{G(q)(1-e^{-\tilde{\lambda}})}\left[\sum_{n=0}^{\infty}\frac{e^{-\tilde{\lambda}}\tilde{\lambda}^n}{n!}G(q)^n - e^{-\tilde{\lambda}}\right] = \frac{F(q)-e^{-\tilde{\lambda}}}{G(q)(1-e^{-\tilde{\lambda}})}$$

Among techniques that use i as a supplier, the fraction that deliver efficiency less than q is $H\left(\frac{q}{q_i^\alpha}\right)$, with density $\frac{1}{q_i^\alpha}H'\left(\frac{q}{q_i^\alpha}\right)$. We can now characterize the probability that a particular technique is the potential buyer's best technique:

$$\Pr\left(\phi \text{ is used}|q_i\right) = \int_0^\infty \frac{1}{q_i^\alpha}H'\left(\frac{\tilde{q}}{q_i^\alpha}\right)\frac{F(\tilde{q})-e^{-\tilde{\lambda}}}{G(\tilde{q})(1-e^{-\tilde{\lambda}})}d\tilde{q}$$

How many known techniques use a given firm as a potential supplier? In other words, how many potential customers does a given supplier have? Across all firms, the distribution over the number of potential customers follows a Poisson law with mean $\tilde{\lambda}$. Each one of those techniques has an equal chance of being the potential buyer's best technique, so the distribution over the number of *actual* customers will also be a Poisson, with parameter:

$$\tilde{\lambda}\int_0^\infty \frac{1}{q_i^\alpha}H'\left(\frac{\tilde{q}}{q_i^\alpha}\right)\frac{F(\tilde{q})-e^{-\tilde{\lambda}}}{G(\tilde{q})(1-e^{-\tilde{\lambda}})}d\tilde{q}$$

Using the functional form for H and taking the limit as $z_0 \to 0$ yields

$$\lambda\frac{q_i^{\alpha\zeta}}{\theta}$$

So among firms with efficiency q, the distribution over the number of customers is a Poisson distribution with parameter $\lambda\frac{q^{\alpha\zeta}}{\theta}$. One can see that the distribution among high efficiency suppliers first order stochastically dominates the distribution among low efficiency suppliers: high efficiency firms get more customers.

Figure 2 shows the average number of customers at each percentile in the efficiency grouping for different values of α. As expected, higher efficiency firms are able to attract more customers than low efficiency firms.

It is notable that curve in Figure 2 depends on only one parameter, α (the share of inputs in production). This is because buyers choose to use the technique that gives the best combination of efficiency (z) and input cost (q). α determines the relative importance of these two factors. Recall that the efficiency associated with a single technique is zq^{α}. If α is large then the the share of inputs is higher, and the cost of inputs becomes relatively more important. An increase in α makes the techniques using high efficiency suppliers even more cost effective. In contrast, when α is low, more weight is put on the idiosyncratic productivity associated with the technique. Because the productivity draws are drawn from the same distribution regardless of the efficiency of the supplier, lowering α increases the odds that a low efficiency firm will be able to attract customers; the low efficiency becomes less relevant to its customers.

When α is higher, more weight is put on the cost of inputs, so the advantage of the high efficiency suppliers is even larger. The shift is evident in Figure 2: with higher α, the high efficiency firms are capturing a much larger share of the customers, at the expense of the low efficiency firms.[17]

We next look at the unconditional distribution over the number of customers among all firms. To find the mass of suppliers with n customers, we simply integrate over suppliers of each efficiency. The resulting formula is given in the following proposition, which also describes the tail of the

[17] It is easy to show a single crossing property: Curves for different values of α in Figure 2 cross exactly once.

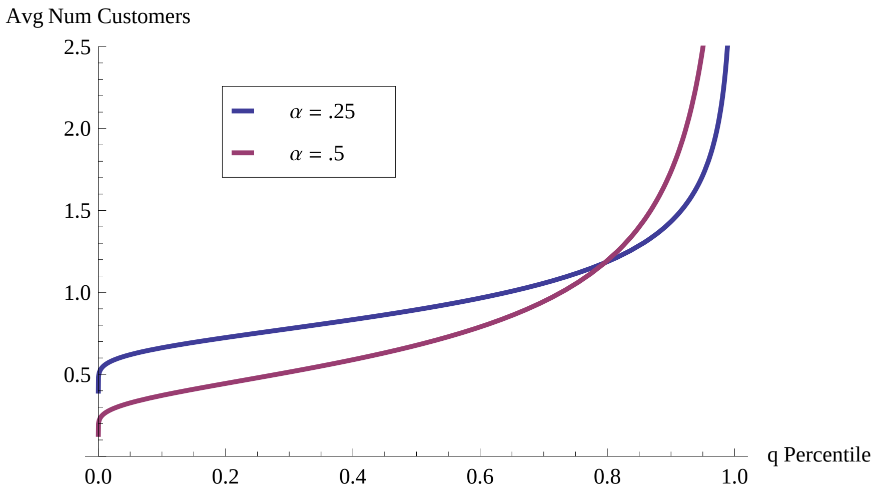


Figure 2: Conditional Distribution of Customers

The mean number of actual customers for each percentile in the efficiency distribution.

distribution.

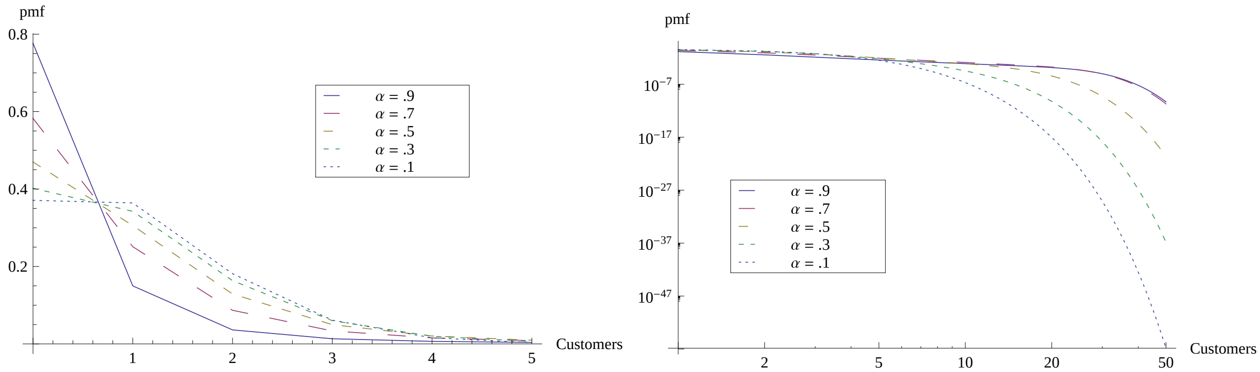
Proposition 2 *Let p_n be the mass of firms with n customers. Then*

$$p_n = \int_0^\infty \frac{\left(\frac{w^{-\alpha}}{\Gamma(1-\alpha)}\right)^n e^{-\left(\frac{w^{-\alpha}}{\Gamma(1-\alpha)}\right)}}{n!} e^{-w} dw$$

The counter cumulative distribution has a tail index of $1/\alpha$.

It is notable that the distribution of number of customers has only a single parameter, α. In particular the density of the network, λ does not enter this equation. While a higher λ raises the number of potential customers for any firm, it also raises the number of alternatives available to each of those potential customers.

Figure 3 shows the distribution of customers for different values of α. When α is high, more weight is put on the cost of inputs, so the distribution is more skewed. The tail is thicker, and but there are also more firms without any customers. In contrast, when α is low, the middle of the



(a) Number of Customers

(b) Number of Customers, Log-Log

Figure 3: The Distribution of Customers

Figure 3a gives the mass of firms with n customers for several different values of the input share α. Figure 3b also gives the mass of firms with n customers, but on a log-log plot, to better show the tail of the distribution.

distribution is thicker.

2.2 Comparison to Preferential Attachment

The preferential attachment model of Barabasi and Albert (1999) was designed to match the scale free nature of the distribution of links in several real world networks. For example, relative to a network in which links are formed randomly and with uniform probability, there are too many websites with many incoming weblinks and also too many with very few incoming weblinks. Roughly, the preferential attachment model explains this as follows: There is an initial network and over time new links are formed and new firms are born. The probability that a new link involves a particular firm is increasing in the number of links that firm already has.

This paper gives an alternative explanation for skewed cross-sectional distribution of input-output links. Because entrepreneurs differ in productivity (some have more efficient supply chains than others), they will also differ in their ability to attract customers. While the distribution of *potential* customers is the same as one would expect in a uniformly random network, the distribution of *actual* customers is not. This is precisely because of endogenous selection: some of those firms with many potential customers are able to offer lower prices and win over a larger fraction of those potential customers.[18]

2.3 Distribution of Employment

In characterizing the cross sectional distribution of employment, we first derive a convenient fact. Let $B\left(q|q_i\right)$ be the the CDF of the efficiency of customers of suppliers with efficiency q_i. In other words, $B\left(q|q_i\right)$ is the distribution efficiency among firms whose best technique uses a supplier with efficiency q_i. We will show that in the limit as $z_0 \to 0$, $B\left(q|q_i\right) = F(q)$. We can solve for this distribution of customers' actual efficiency with an application of Bayes rule:

$$B'\left(q|q_i\right) = \frac{\frac{1}{q_i^\alpha}H'\left(\frac{q}{q_i^\alpha}\right)\frac{F(q)-e^{-\tilde{\lambda}}}{G(q)(1-e^{-\tilde{\lambda}})}}{\int_0^\infty \frac{1}{q_i^\alpha}H'\left(\frac{\tilde{q}}{q_i^\alpha}\right)\frac{F(\tilde{q})-e^{-\tilde{\lambda}}}{G(\tilde{q})(1-e^{-\tilde{\lambda}})}d\tilde{q}}$$

The numerator is the density of efficiency delivered by techniques that use suppliers with q_i, $\frac{1}{q_i^\alpha}H'\left(\frac{q}{q_i^\alpha}\right)$, multiplied by the probability that a such a technique is the potential customer's best technique. The denominator is the probability that a technique from i is the customer's best technique. We can use the functional forms for H and take a limit as $z_0 \to 0$:

$$\begin{aligned}
\lim_{z_0\to 0} B'\left(q|q_i\right) &= \lim_{z_0\to 0}\frac{\zeta z_0^\zeta q_i^{\alpha\zeta} q^{-\zeta-1}e^{-\theta q^{-\zeta}}}{\int_{z_0 q_i^\alpha}^\infty \zeta z_0^\zeta q_i^{\alpha\zeta}\tilde{q}^{-\zeta-1}e^{-\theta\tilde{q}^{-\zeta}}d\tilde{q}} \\
&= \zeta\theta q^{-\zeta-1}e^{-\theta q^{-\zeta}} \\
&= \lim_{z_0\to 0} F'(q)
\end{aligned}$$

[18] There is an analogous comparison for the world wide web. The preferential attachment model says the www.nytimes.com has many incoming links because it happened to have many such links initially, and consequently many other websites linked to it as well. In contrast, a model with endogenous network formation would give the explanation that www.nytimes.com offers better content than other sites, and therefore other sites choose to link to it more frequently.

This implies that knowing the efficiency of a supplier gives no information about the identity of its customer. That is, there are no systematic differences between the customers of low and high efficiency suppliers.

This has two implications. First, it gives insight into the determinants of size. Aside from more sales of final goods, high efficiency firms will on average be larger because they have more customers, not because their customers are any bigger.

Second we can treat the characteristics of customers as independent, identically distributed random variables. This will be helpful in several ways. Of particular use here is the fact that we can treat the size of a customer as an IID random variable.

Distribution of Employment

We are interested in adding together the labor used to make goods for final consumption and for intermediate use for each customer. Since the latter can be treated as independent random variables, it is easiest to work with characteristic functions of the relevant distributions.

Let $\chi(s)$ be the characteristic function associated with the cross sectional distribution of employment. To get at it, we will create a fixed point problem. Roughly, the information that is used includes:

- Given a firm's efficiency, the quantity of labor used for production of final output.
- Given a firm's efficiency, the distribution over the number of actual customers.
- The fact that the quantity of labor used by a customer is independent of the efficiency of the supplier.
- The fact that if a customer uses L units of labor, the supplier will use αL units of labor to make the inputs for that customer.

As described in Appendix D.2, the characteristic function solves the equation:

$$\chi(s) = \int_0^\infty \exp\left\{ is \frac{t^{-\frac{\varepsilon-1}{\zeta}}}{\Gamma\left(1-\frac{\varepsilon-1}{\zeta}\right)} (1-\alpha) L - \frac{t^{-\alpha}}{\Gamma(1-\alpha)} [1-\chi(\alpha s)] - t \right\} dt \tag{13}$$

There are several things to note. First, as L changes, the distribution of labor scales uniformly.[19] Second the distribution depends on only two parameters, α and $\frac{\varepsilon-1}{\zeta}$. The share of intermediates matters for the same reason as before, it determines the skew of the distribution of number of customers. $\frac{\varepsilon-1}{\zeta}$ is a composite of two parameters, the elasticity of substitution in final consumption, and ζ, the tail index of both $H(\cdot)$ (the Pareto distribution from which productivity shocks are drawn) and $F(\cdot)$ (the cross sectional distribution of efficiency). In combination, these parameters determine the tail index of the distribution of final consumption. When ζ is small, the efficiency distribution has a thicker tail, inducing a thicker tail in the distribution of final consumption. When ε is high, consumers are more willing to substitute toward low cost goods, also thickening the tail of final consumption.

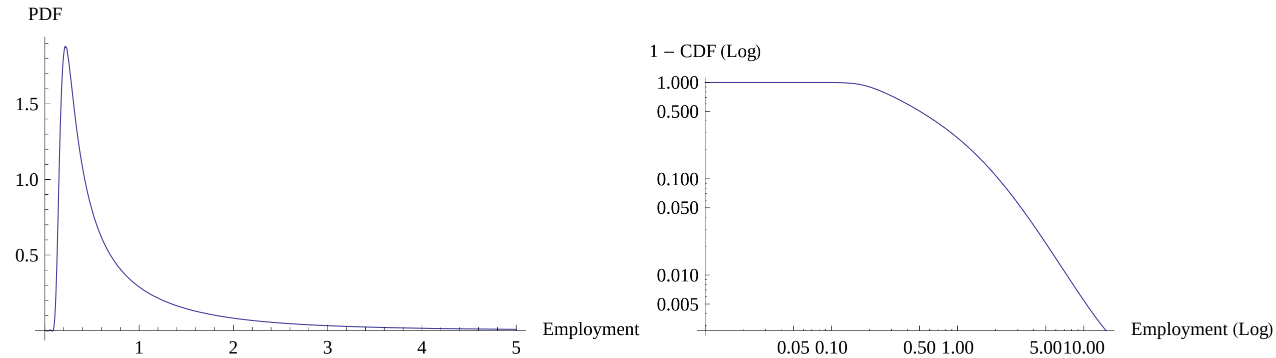
Equation (13) can be used to solve for $\chi(s)$ numerically.[20] We can consequently use standard methods to back out the distribution of employment form its characteristic function.

Figure 4 shows the distribution for the parameters $\alpha = \frac{\varepsilon-1}{\zeta} = 1/2$ and $L = 1$. One can see that this density is quite skewed, with the mode well below the mean of 1.

[19] In fact, one could write the characteristic function associated with the fraction of labor used by each firm as $\hat{\chi}(s) = \chi\left(\frac{1}{L}s\right)$. This satisfies

$$\hat{\chi}(s) = \int_0^\infty \exp\left\{ is \frac{t^{-\frac{\varepsilon-1}{\zeta}}}{\Gamma\left(1-\frac{\varepsilon-1}{\zeta}\right)} (1-\alpha) - \frac{t^{-\alpha}}{\Gamma(1-\alpha)} [1-\hat{\chi}(\alpha s)] - t \right\} dt$$

[20] While equation (13) is a functional equation, it resembles a difference equation. One can solve this using a reverse shooting algorithm, starting near the point $\chi(0) = 1$ and interpolating.



(a) Density of Employment

(b) Right CDF of Employment, Log-Log

Figure 4: The Distribution of Employment

Figure 4a gives the density of employment with $\alpha = \frac{\varepsilon - 1}{\zeta} = 1/2$. Figure 4b gives right CDF on a log-log plot to show the shape of the tail of the distribution.

3 Market Structure and Decentralized Equilibria

We will consider a market structure in which there is monopolistic competition across final goods but bilateral two part tariffs for intermediate goods. In addition we will consider contracting terms that are "pairwise stable."[21] This equilibrium concept is reminiscent of Nash equilibrium, except there is no explicit description of a formal game. We will characterize the set of contracting arrangements for which the are no profitable unilateral or mutually beneficial pairwise deviations.

As an equilibrium concept, pairwise stability captures the dynamic spirit of the model. As new techniques are found, some of the cost savings are passed on to downstream customers in the form of lower prices. An important aspect of this is that when the terms of a contract are no longer optimal, those terms can be renegotiated to reflect the change in cost structure in the supply chain. While the equilibrium definition is a static one, the motivation is inherently dynamic, and pairwise

[21] As an equilibrium concept, pairwise stability is used frequently in the networks literature (Jackson (2008) provides an excellent survey). For many applications in this literature, payoffs are a function of the particular links that are formed, and the idea is to find a network of links for which no pair of nodes wish to change whether or not they are connected (see Jackson (2003)). Here, the usual concept is extended to include contractual terms of trade in any input output relationship.

deviations capture the spirit of the renegotiation that would take place over time.[22]

One result will be that for each technique that is used in equilibrium, marginal price equals marginal cost. While the space of possible contracts is fairly large, it seems reasonable to focus on a space of contracts that allows for this possibility. Anything other than this kind of two part tariff would lead to double marginalization and leave surplus on the table. Input-output relationships are generally long-lived, so it would be surprising if the contracting terms remained inefficient. In addition, the informational demands of these contracts (in equilibrium) are not large. Firms do not need to know that much about the environment to get the terms of the contract right.[23] While these concerns are outside the model, they motivate the focus on two part pricing.

Monopolistic competition across final goods leads to a uniform markup of $\frac{\varepsilon}{\varepsilon-1}$. This gives surplus for each final good that is divided across firms in that supply chain. We will show that allowing for two part tariffs decentralizes the planner's solution: in this decentralization, the inputs, outputs, and labor used by each firm are the same as the the efficient allocation. Each firm will choose the technique that gives the best combination of cost/productivity. There are many ways to divide the surplus from sales of final goods across supply chains, but this is not relevant for the allocation of resources. Since the marginal input price equals marginal cost, the quantity supplied will be efficient, and since labor is supplied inelastically the monopoly markup on final goods is not

[22] Another a solution concept, Nash bargaining, is tricky precisely because of renegotiation: there is no well defined outside option. If a supplier i and buyer j are bargaining over the contract terms for a technique, it is natural that if the bargaining breaks down j will use its next best supplier. However, if this happens, does j renegotiate with its current contracts with its customers? More generally, the Nash bargaining solution requires well defined outside options, and to do this would require a well defined order of negotiation and renegotiation across pairs of firms.

[23] Another consideration is that allowing for bilateral two part tariffs leads to a characterization of equilibrium that is tractable. Consider an alternative in which firms set a single price for all customers. A difficulty in solving for the optimal prices in this world is that demand curves facing firms are not continuous let alone differentiable, as lowering a price a little may allow a supplier to beat out a competitor and give a spike in quantity demanded (or may allow the buyer to lower *its* price enough to beat out *its* competitor, giving that buyer and consequently the supplier a spike in quantity demanded). If two part tariffs are not available, solving for the optimal prices (and consequently the allocation) is quite challenging.

distortionary.[24]

3.1 Pairwise Stable Equilibrium

Here we define a pairwise stable equilibrium. This has two stages. First an "arrangement" determines which techniques are used and pricing for each of those techniques. Second, firms set prices of final goods and select their input mix. In defining an arrangement, care is taken to ensure enough structure so that payoffs after deviations are well defined.

Definition 1 *An* ***arrangement*** *consists of the following:*

(i) For each technique $\phi \in \Phi(t)$*, define the triple* $\{p(\phi), \tau(\phi), A(\phi)\} \in \mathbb{R}^{+} \times \mathbb{R} \times \{0,1\}$*.*

$p(\phi)$ *and* $\tau(\phi)$ *are the price and fixed part of a two part tariff.* $A(\phi) = 1$ *implies that the supplier associated with the technique,* $s(\phi)$*, is willing to supply goods at those terms, whereas* $A(\phi) = 0$ *indicates she is not.*

(ii) For each firm, j*, a ordering of techniques. Formally, this is a bijective function* $O_j : \Phi_j^S \rightarrow \left\{1, ..., \left|\Phi_j^B\right|\right\}$*.*

Let $\phi_j^* \equiv \arg\min_{\left\{\phi \in \Phi_j^S | A(\phi)=1\right\}} O_j(\phi)$ be the technique actually used by firm j in the arrangement. This is jointly determined by the ordering O_j and the whether or not each of j's potential suppliers are willing to accept the contracting terms of the arrangement. $\Phi_j^{B*} \equiv \left\{\phi \in \Phi_j^B | \phi = \phi_{b(\phi)}^*\right\}$ is the set of customers that use firm j's good as an input in the arrangement.

Given an arrangement, firms choose a price and quantity for final output along with an input mix to maximize profit. $p_j^0 y_j^0$ is revenue from final output, $\sum_{\phi \in \Phi_j^{B*}} [p(\phi)x(\phi) + \tau(\phi)]$ is revenue from output for use as an intermediate input, and $\tau\left(\phi_j^*\right) + p\left(\phi_j^*\right) x\left(\phi_j^*\right)$ is the cost of intermediates

[24] If labor were supplied elastically, the only differences between the planner's allocation and the decentralized equilibrium would be that less labor would be supplied because of the monopoly markups and all production would scale down in proportion to the decrease in aggregate labor.

inputs used in production. Profit for firm j is

$$\pi_j = \max_{p_j^0, y_j^0, L, x} p_j^0 y_j^0 + \sum_{\phi \in \Phi_j^{B*}} [p(\phi)x(\phi) + \tau(\phi)] - \tau\left(\phi_j^*\right) - p\left(\phi_j^*\right) x - wL$$

subject to the technological constraints and a downward sloping demand curve for final output:

$$\begin{aligned} y_j^0 + \sum_{\phi \in \Phi_j^{B*}} x(\phi) &\leq \frac{1}{\alpha^\alpha (1-\alpha)^{1-\alpha}} z\left(\phi_j^B\right) x\left(\phi_j^B\right)^\alpha L^{1-\alpha} \\ y_j^0 &\leq D\left(p_j^0\right) \end{aligned}$$

Given an arrangement, $\left\{\{p(\phi), \tau(\phi), A(\phi)\}_{\phi \in \Phi_j^B}, O_j\right\}_{j \in J}$, consider two types of deviations:

A **unilateral deviation** for firm j is an alternative order $\hat{O}_j$ and/or an alternative acceptance $\hat{A}(\phi)$ for each $\phi \in \Phi_j^B$.

For a technique ϕ, a **pairwise deviation** is an alternative two part tariff $\hat{p}(\phi), \hat{\tau}(\phi)$, an alternative acceptance for the supplier $s(\phi)$ about whether to accept the contract, $\hat{A}(\phi)$, and an alternative ordering for the buyer, $\hat{O}_{b(\phi)}$.

Either of these types of deviations delivers an alternative arrangement, and hence alternative payoffs.

Definition 2 *A pairwise stable equilibrium is an arrangement* $\left\{\{p(\phi), \tau(\phi), A(\phi)\}_{\phi \in \Phi_j^B}, O_j\right\}_{j \in J}$, *firms choices,* $\left\{p_j^0, y_j^0, x_j, L_j\right\}_{j \in J}$, *and a wage* w *such that*

- *Given wages and profits, the final consumers maximize utility.*

- *For each* $j \in J$, $\left\{p_j^0, y_j^0, x_j, L_j\right\}$ *maximize firm* j*'s profit given the arrangement, wages, and the final demand.*

- *Labor and final goods markets clear*
- *There are no unilateral deviations that would increase a firm's profit.*
- *There are no pairwise deviations that would increase each firms' profit.*

Let q_j^{pw} be the efficiency of firm j, equal to the wage divided by marginal cost. In Appendix E we prove the following:

Proposition 3 *In any pairwise stable equilibrium:*

(i) For each technique that is actually used, price is equal to marginal cost, $p\left(\phi_j^*\right) = \frac{w}{q_j^{pw}}$*, and the fixed fee is nonnegative,* $\tau\left(\phi_j^*\right) \geq 0$*;*

(ii) $q_j^{pw} = \max_{\phi \in \Phi_j^S} z(\phi)\left(q_{s(\phi)}^{pw}\right)^{\alpha}$

These conditions are the same as the necessary conditions to the planners problem. In fact, there are many pairwise stable equilibria that decentralize the planner's solution. In each of these, marginal costs are the same as in the planner's solution, but the various equilibria differ in the fixed part of the tariffs. In each of these equilibria, the markup in the sales of goods to the final consumer generates a profit for each good. This profit is then divided up across the supply chain to produce that good according to the fixed fees.[25] Normally the markups would distort the consumption-leisure margin, but since labor is supplied inelastically, the markup has no effect on the allocation of goods and labor. Even if labor were elastic, all of these equilibria would generate the *same aggregate production function* as the planner would choose.

There are additional equilibria which do not decentralize the planner's problem. For example, there is always an equilibrium in which all firms set a price of infinity (this corresponds to one of

[25] There is an upper limit to the fees that can be charged in equilibrium: they cannot be so high that the buyer would find it beneficial to use an alternative supplier. Short of that limit, any nonnegative fee will do.

constant fixed points of equation (10)).

There is a sense in which that equilibrium is fragile (along with any others that might exist). There is a slightly perturbed economic environment no such equilibrium exists. More specifically, consider adding to each firm's techniques a production function that uses only labor, $y = \underline{q}L$, with $\underline{q}$ common to all firms. In this environment we can provide a welfare theorem: every pairwise stable equilibrium decentralizes the planner's solution.

Proposition 4 *If $\underline{q} > 0$, then with probability one $F^{sp} = F^{pw}$ and $Q^{sp} = Q^{pw}$.*

See Appendix E for proof. For the remainder of the paper, we will focus on equilibria that decentralize the planner's allocation.

3.2 An Alternative Interpretation of the Model

An alternative interpretation of the model is that the technology embodied in a technique is non-rival and freely available for others to replicate. In this interpretation, each good is produced by an island of identical firms, and labor is perfectly mobile across islands. The state of technology can still be represented as a network, but each node is an island of firms producing a single good rather than an individual entrepreneur. Again, the network represents the input-output architecture of the economy, but among islands of firms. In this interpretation perfect competition might be a more natural benchmark with all prices are set at marginal cost. It is straightforward to show that competitive equilibrium allocations correspond to those of pairwise stable equilibria.

4 Asymmetric Networks

The preceding analysis studied a very specific type of network, and leaves open the question of how alternative network configurations would affect aggregate productivity. For example, if many firms are in the center of a city and others are in the outskirts, how would aggregate productivity respond to increased concentration? If some entrepreneurs were particularly active in finding new techniques (and others particularly inactive), how would this change patterns of diffusion of productivity gains from newly discovered techniques?

To get at these, we first extend the previous setup to allow for more interesting network configurations. There are multiple types of firms, indexed by $k \in K$. The only structural difference between firms of different types is how frequently they are involved with new techniques that are discovered. Let M_t be the mass of each type k firms and, abusing notation, let K be the number of types (in addition to the set of types).

The social planner's problem is exactly the same as in Section 1.2 and all first order conditions carry over. Instead of characterizing the distribution of efficiency across all firms, it will be convenient to characterize the distribution among firms of each type. Let $F_k(q)$ be the fraction of type k firms with efficiency less than q. We proceed to characterize these distributions by setting up a fixed point problem.

We next define several objects that have analogs in Section 1. $\omega_k(n; k')$ is the fraction type k firms that have access to n techniques with suppliers of type k'. At a given point in time this follows a Poisson distribution with mean $\tilde{\lambda}_k(k')$. Let $G_k(q; k')$ be the distribution of efficiency provided by a single technique drawn by a type k firm with supplier type k'. $G_k(q;k')$ is then:

$$G_k(q;k') = \int_0^\infty F_{k'}\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right) dH(z)$$

Given $G_k(q;k')$, we can compute $\Pr\left(q_j \le q | j \text{ is type } k\right)$, the probability that, given all of its draws of techniques, a firm of type k has efficiency no greater than q:

$$\Pr\left(q_j \le q | j \text{ is type } k\right) = \prod_{k' \in K} \sum_{n=0}^{\infty} \omega_k(n;k') G_k(q;k')^n = e^{-\sum_{k'} \tilde{\lambda}_k(k')[1-G_k(q;k')]}$$

The same abuse of the law of large numbers gives $F_k(q) = \Pr\left(q_j \le q | j \text{ is type } k\right)$, giving the fixed point problem, K functional equations for the K unknown functions $\{F_k(\cdot)\}_{k \in K}$:

$$F_k(q) = e^{-\sum_{k' \in K} \tilde{\lambda}_k(k')\left[1 - \int_0^\infty F_{k'}\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right) dH(z)\right]} \tag{14}$$

4.1 A Parametric Assumption

We will use the same functional forms as in the one type model, $H(z) = 1 - \left(\frac{z}{z_0}\right)^{-\zeta}$ and $\tilde{\lambda}_k(k') = \lambda_k(k') z_0^{-\zeta}$. We will then look at the equilibrium of the limiting economy as $z_0 \to 0$. With a similar argument, we will show that any set of solutions to equation (14) must follow Frechet distributions. As before we can write

$$\tilde{\lambda}_k(k')\left[1 - G_k(q;k')\right] = q^{-\zeta} \lambda_k(k') \int_0^{\left(\frac{q}{z_0}\right)^{\frac{1}{\alpha}}} \alpha\zeta x^{\alpha\zeta - 1}\left(1 - F_{k'}(x)\right) dx$$

Substituting into equation (14) for each k' we get

$$-\log F_k(q) = q^{-\zeta} \sum_{k' \in K} \lambda_k(k') \int_0^{\left(\frac{q}{z_0}\right)^{\frac{1}{\alpha}}} \alpha\zeta x^{\alpha\zeta-1} \left(1 - F_{k'}(x)\right) dx$$

For any set of $\{F_k(\cdot)\}_{k\in K}$, as $z_0 \to 0$, this expression goes to $q^{-\zeta}$ multiplied by a constant. For each k, label this constant θ_k so that $F_k(q) = e^{-\theta_k q^{-\zeta}}$.

We next solve for $\{\theta_k\}_{k\in K}$, which are defined to satisfy

$$\theta_k = \sum_{k' \in K} \lambda_k(k') \int_0^{\infty} \alpha\zeta x^{\alpha\zeta-1} \left(1 - F_{k'}(x)\right) dx = \sum_{k' \in K} \lambda_k(k') \int_0^{\infty} x^{\alpha\zeta} F'_{k'}(x) dx$$

Plugging in the functional form for $F_{k'}(q)$ gives

$$\theta_k = \Gamma(1-\alpha) \sum_{k' \in K} \lambda_k(k') \theta_{k'}^{\alpha} \tag{15}$$

Notice also that, as before, for any $\{\lambda_k(k')\}_{k,k' \in K}$ there are three solutions to equation (15): $\theta_k = 0, \forall k \in K$, $\theta_k = \infty, \forall k \in K$, and a third solution that is the solution to the planner's problem.

4.2 Aggregate Output

Given the distribution of efficiency across firms, total output will be $Y^0 = QL$ where $Q^{\varepsilon-1} = \int_J q_j^{\varepsilon-1}$ (again, the analysis in the one type economy carries over). It will be convenient to define $Q_k \equiv \left(\int_0^\infty q^{\varepsilon-1} dF_k(q)\right)^{\frac{1}{\varepsilon-1}}$ to be a productivity aggregator among firms of type k. We can then write

$$Q^{\varepsilon-1} = \sum_{k\in K} M_k Q_k^{\varepsilon-1}$$

With the functional forms, we can write (as before):

$$Q_k = \theta_k^{1/\zeta}\Gamma\left(1 - \frac{\varepsilon - 1}{\zeta}\right)^{\frac{1}{\varepsilon-1}}$$

The productivity aggregator for the whole economy can then be written as

$$Q^{\varepsilon-1} = \Gamma\left(1 - \frac{\varepsilon - 1}{\zeta}\right)\sum_{k\in K} M_k\theta_k^{\frac{\varepsilon-1}{\zeta}}$$

5 Superstars and Productivity Spillovers

The purpose of this section is to demonstrate how changing the configuration of the network affects aggregate output. We consider examples in which there are two types of firms, indexed by $k \in \{A, B\}$. The total mass of firms is $M_A + M_B = 1$. Again, the types of firms differ only in how frequently they are involved with new techniques that are discovered. For a firm of type k let λ_k be the (cumulative) arrival rate of techniques, so that $\lambda_k = \sum_{k'} \lambda_k(k')$.

To focus on the influence of the configuration, I will hold the total number of techniques in the network constant but vary their distribution. In other words, I will let $\{\lambda_k(k')\}$ and $\{M_k\}$ vary subject to

$$\lambda = M_A\left(\lambda_A(A) + \lambda_A(B)\right) + M_B\left(\lambda_B(A) + \lambda_B(B)\right)$$

holding λ constant.

The heterogeneity across types is parameterized as follows: Let $\rho_s \equiv \frac{\lambda_A}{\lambda_B} = \frac{\lambda_A(A)+\lambda_A(B)}{\lambda_B(A)+\lambda_B(B)}$ (s is for "supplier"). This is a measure of how much more frequently type A firms discover new techniques. When ρ_s is larger, type A has a larger advantage in finding potential suppliers.

Similarly, let $\rho_b \equiv \frac{(\lambda_A(A)+\lambda_B(A))/M_A}{(\lambda_A(B)+\lambda_B(B))/M_B}$ (b is for buyer). This is a measure of how much more frequently type A firms find potential buyers. When ρ_b is larger, type A has a larger advantage in finding potential buyers. If $\rho_b = 1$ then the probability of being a supplier is uniform across all firms. If $\rho_b > 1$, then a new technique is relatively more likely to use a type A firm than would be suggested by M_A and M_B.

Given the values of ρ_s and ρ_b (along with the assumption that $\frac{\lambda_A(B)}{\lambda_A(A)} = \frac{\lambda_B(B)}{\lambda_B(A)}$ which can easily be abandoned) we can solve for the implied values of $\{\lambda_A(A), \lambda_A(B), \lambda_B(A), \lambda_B(B)\}$:

$$\begin{aligned}
\lambda_A(A) &= \frac{\rho_b M_A}{\rho_b M_A + M_B} \frac{\rho_s}{\rho_s M_A + M_B} \lambda \\
\lambda_A(B) &= \frac{M_B}{\rho_b M_A + M_B} \frac{\rho_s}{\rho_s M_A + M_B} \lambda \\
\lambda_B(A) &= \frac{\rho_b M_A}{\rho_b M_A + M_B} \frac{1}{\rho_s M_A + M_B} \lambda \\
\lambda_B(B) &= \frac{M_A}{\rho_b M_A + M_B} \frac{1}{\rho_s M_A + M_B} \lambda
\end{aligned}$$

To compute aggregate output, we need an expression for the productivity aggregator $Q = \left(\int_J q_j^{\varepsilon-1}\right)^{\frac{1}{\varepsilon-1}}$. Given the values of $\{\lambda_k(k')\}_{k,k' \in K}$ we can solve for the values of θ_A and θ_B using equation (15). With this, we can then solve for the productivity aggregator Q:

$$Q^{\varepsilon-1} = M_A Q_A^{\varepsilon-1} + M_B Q_B^{\varepsilon-1} \propto M_A \theta_A^{\frac{\varepsilon-1}{\zeta}} + M_B \theta_B^{\frac{\varepsilon-1}{\zeta}}$$

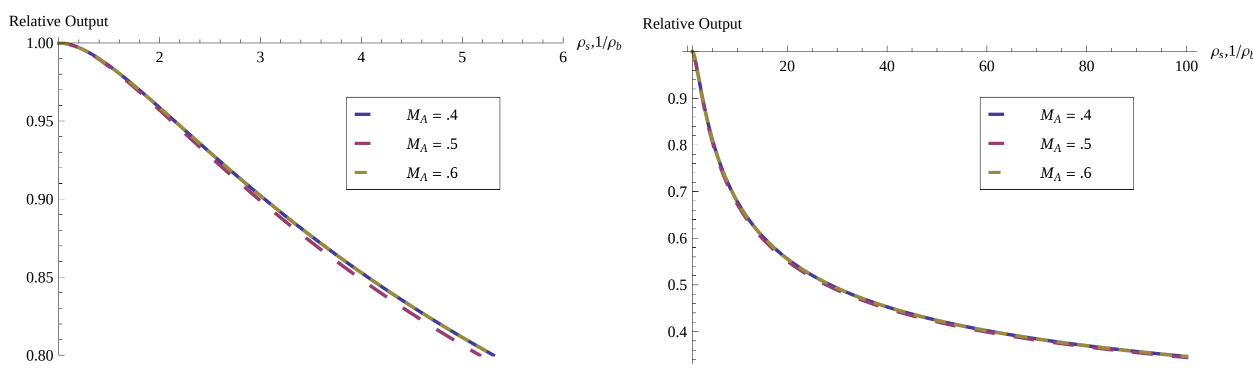
where Q_A and Q_B are the productivity aggregators among firms of each type. We are interested in how Q varies as the structure of the network changes.

Specialization

We first consider an environment in which firms specialize in activities. Type A firms have an advantage in finding new production techniques ($\rho_s > 1$) but type B firms have an advantage in finding potential buyers ($\rho_b < 1$). One can think of type A firms as specializing in R&D and type B firms as specializing in marketing.

Figure 5 shows aggregate output relative to a uniform network ($\rho_s = \rho_b = 1$) with the same number of total techniques. It is evident that when firms specialize, Aggregate output is quite a bit lower than a uniformly random network.

Because the R&D firms discover more techniques, they are more productive. However, the firms specializing in marketing are acquiring a larger share of potential customers. The cost savings found by the R&D firms get less of a chance to be passed on to other firms: the wrong firms are supplying the inputs.



(a) Small Scale (b) Large Scale

Figure 5: Specialization: Varying $\rho_s > 1$, $\rho_b < 1$

Figure 5 shows aggregate output relative to a uniform network for various values of activity (Figure 5a and Figure 5b show the same graph at different scales). When ρ_s is larger, type A firms discover techniques more frequently. When $\rho_b < 1$, type B firms find potential buyers more frequently.

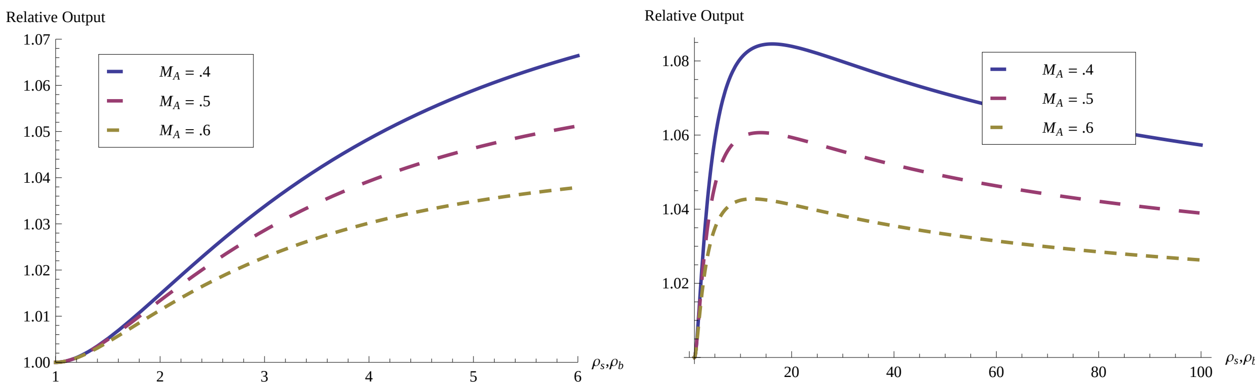
Hubs

We first consider the case in which type A firms are hubs: they are more likely to discover new techniques and they are relatively more likely to be the supplier when other firms get new techniques.

This scenario is motivated by a partially urbanized economy in which urban entrepreneurs interact with each other frequently, while rural entrepreneurs are less active. Formally we set $\rho_s = \rho_b$ and vary this common number for various values of M_A. Figure 6 shows aggregate output relative to a uniform network ($\rho_s = \rho_b = 1$) with the same number of techniques per firm. When the network is close to uniform, aggregate output rises with the increased concentration of techniques among type A firms. However this relationship is non-monotonic; as $\rho_s, \rho_b \to \infty$, relative productivity slowly falls back to 1.

Aggregate output is higher than in the uniform network because the more productive firms are becoming the important suppliers in the economy. When $\rho_s > 1$, type A firms draw many techniques and are therefore more likely to have high efficiency. While type B firms draw a smaller number of techniques, the techniques they do draw are likely to have type A firms as suppliers, meaning that the type B firms are likely to end up with a low marginal cost. When $\rho_b, \rho_s \to \infty$, the impact disappears, because all type B firms are essentially disconnected from the network and are no longer able to benefit from the high efficiency of the type A firms. Type A firms are extremely productive (there are more techniques among them) but fewer firms are producing, and consequently less gains from variety. These effects exactly offset, and aggregate output is the same as the uniform case.

For intermediate values, the increase in aggregate output in response to concentration is larger when M_A is small. This happens because more techniques are concentrated within the type A firms, so productivity among those firms is high. Type B firms are likely to have techniques that use type A firms, so that they are increasingly able to benefit from the high productivity among type A that is due to the increased concentration. This is most stark as $M_A \to 0$, in which case the peak of the curve rises unboundedly. More formally, for a given M_A, set $\rho_b = \rho_s = \frac{1/2}{M_A}$. Then



(a) Small Scale

(b) Large Scale

Figure 6: Type A are Hubs: Varying Both ρ_s and ρ_b

Figure 6 shows aggregate output relative to a uniform network for various values of activity (Figure 6a and Figure 6b show the same graph at different scales). When the common value of ρ_p and ρ_s is larger, a larger share techniques involve type A firms.

$\lim_{M_A \to 0} Q = \infty$.[26]

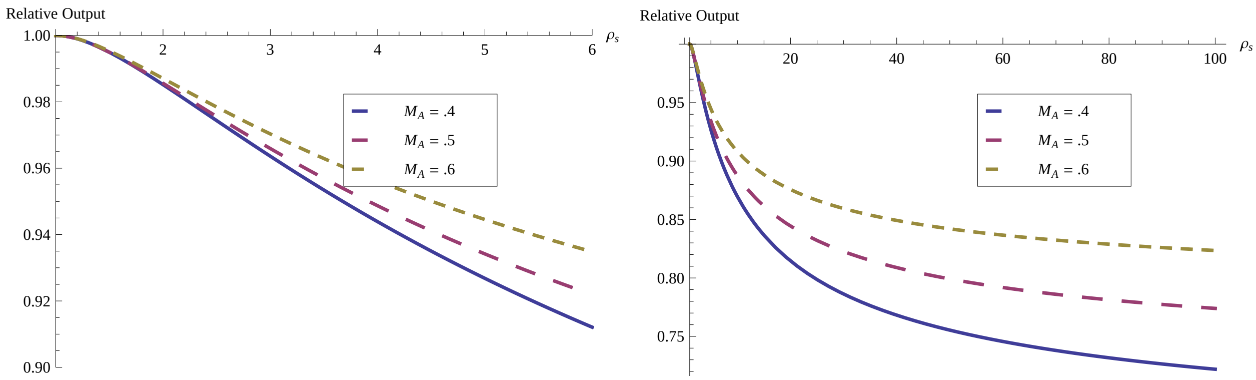
Advantage in Finding Potential Suppliers

We next examine the case in which ρ_s varies but $\rho_b = 1$, shown in Figure 7. In this case type A firms are more likely to discover new techniques, but do not have any advantage in finding potential buyers. Here we can see that total output is smaller than the uniform case (and the drop is persistent: as $\rho_s \to \infty$, relative productivity stays depressed below 1). Here, while type A firms are more productive as a group than they would be in the uniform case, the type B firms are less able to take advantage of this, because the techniques they find aren't especially concentrated on the type A firms. As a result, type B firms are less productive in than they would be in the uniform case, so much so that this dominates the increased productivity among type A firms.

The intuition for why aggregate productivity is lower than the uniform case can be found in

[26] An even more stark example (which is easier to analyze by hand) is $\rho_s = \frac{M_B}{M_A}$ and $\rho_b \to \infty$. This leads to (normalizing $\lambda = 1$), $\lambda_A(A) = \frac{1/2}{M_A}$, $\lambda_B(A) = \frac{1/2}{M_B}$, $\lambda_A(B) = \lambda_B(B) = 0$. Here, no matter how few type A firms there are, half of all total techniques are drawn by those firms. However, *all* techniques use type A firms as inputs. Again, we have $\lim_{M_A \to 0} Q = \infty$.

equation (15) relating the average efficiencies of each type: $\theta_k = \Gamma(1-\alpha)\sum_{k' \in K} \lambda_k(k')\theta_{k'}^{\alpha}$. Note that θ_k is a linear combination of concave functions of each of the $\theta_{k'}$s. Thus θ_k would be higher if the $\theta_{k'}$s were closer together.



(a) Small Scale

(b) Large Scale

Figure 7: Type A discover techniques more frequently: Varying ρ_s with $\rho_b = 1$

Figure 7 shows aggregate output relative to a uniform network for various values of activity (Figure 7a and Figure 7b show the same graph at different scales). When ρ_s is larger, type A firms discover techniques more frequently.

Advantage in Finding Potential Buyers

Lastly we examine the the case in which $\rho_s = 1$ but ρ_b varies, so that A has an advantage in finding potential buyers, but not in finding potential suppliers. As shown in Figure 8, aggregate productivity is exactly the same as in the uniform case. Considering this and the previous case, we can infer the different roles of ρ_b and ρ_s. ρ_s generates productivity differences across the different types, as drawing more techniques leads to (on average) higher productivity. ρ_b determines how much of these productivity differences spill over to the other types. In this case where $\rho_s = 1$, there are no productivity differences to spill over, so varying ρ_b makes no differences. Compare this to $\rho_s \neq 1$, in which case varying ρ_b can make a big difference.

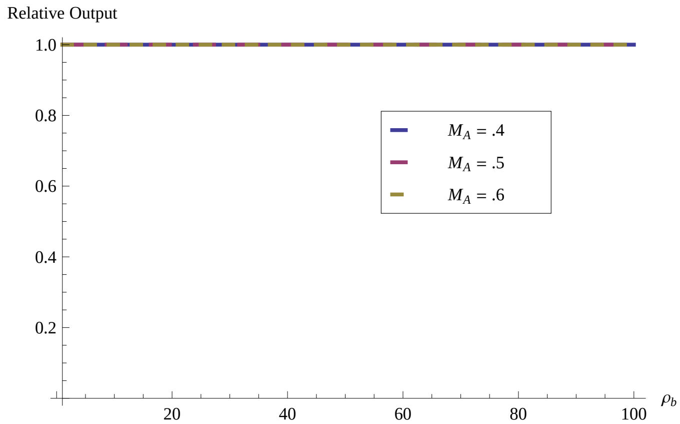


Figure 8: Type A discover potential buyers more frequently: $\rho_s = 1$, Varying ρ_b

Figure 8 shows aggregate output relative to a uniform network for various values of activity. When ρ_b is larger, type A firms are more likely to find potential buyers. This is not a misprint; all three lines take the constant value 1.

6 Multiple Inputs

In the preceding analysis, each firm used a single input in production. While this is assumption makes the analysis (and visualization) of the model economy considerably easier, it is obviously at odds with production in the real world. In this section we extend the analysis so that firms use multiple inputs. While there are several ways to model what happens when firms discover new production methods, I will focus on a version that I consider to be both natural and tractable.

A firm needs to accomplish a fixed set of tasks. Over time, firms discover new techniques to accomplishing each task, and each of these techniques is a sub-production function that requires a particular input. In this way, a firm can improve production by finding techniques that are more cost effective in accomplishing individual tasks.[27]

The central assumption is that each firm uses a production function that combines N tasks into a unit of output. To accomplish each task, the firm uses a technique. For the technique used for

[27] An alternative modeling strategy is that a new technique requires an entirely new input bundle. In that model, a firm would change suppliers all at once, rather than one at a time.

the nth task, let z_n be the cost of inputs and q_n be the efficiency of the supplier associated with that task. Then the production is such that the efficiency of the firm is

$$q = \prod_{n=1}^{N} \left(z_n q_n^{\alpha}\right)^{\frac{1}{N}}$$

The production function that justifies this can be written in two ways that are equivalent as long as the firm minimizes cost:

$$\begin{aligned} y &= \prod_{n=1}^{N} \left(\frac{z_n x_n^{\alpha} L_n^{1-\alpha}}{\left(\frac{1-\alpha}{N}\right)^{1-\alpha} \left(\frac{\alpha}{N}\right)^{\alpha}} \right)^{\frac{1}{N}} \\ y &= \frac{1}{(1-\alpha)^{1-\alpha}} \prod_{n=1}^{N} \left(\frac{\alpha}{N}\right)^{-\frac{\alpha}{N}} \prod_{n=1}^{N} z_n^{\frac{1}{N}} \left(\prod_{n=1}^{N} x_n^{\frac{1}{N}} \right)^{\alpha} L^{1-\alpha} \end{aligned}$$

The difference between the two is that in the first version labor is explicitly allocated separately to each task. The important feature is that α is the share of intermediate inputs.

Let $F(q)$ be the fraction of firms with efficiency no greater than q. We will use the same method as before to solve to define a fixed point problem for F.

For a firm, let v_n be the efficiency of a firm in accomplishing the nth task. Also, the number of techniques a firm has to accomplish a given task is follows a Poisson distribution with mean $\tilde{\lambda}$ (which is the same for each task). Let $K(v)$ be probability that $v_n < v$.[28] Then by the same analysis as before, we can write

$$K(v) = e^{-\tilde{\lambda} \int_0^{\infty} \left[1 - F\left(\left(\frac{v}{z}\right)^{1/\alpha}\right)\right] dH(z)}$$

[28] The analysis can easily be extended so that $\tilde{\lambda}$ varied by task, which would imply that the CDFs of each v_n to be different (i.e. $K_n(\cdot)$).

Since $q = \prod_{n=1}^{N} v_n^{\frac{1}{N}}$, we can write $F(q)$ as

$$F(q) = \int_0^{\infty} ... \int_0^{\infty} K\left(\frac{q^N}{\prod_{n=2}^{N} v_n}\right) dK\left(v_2\right) dK\left(v_3\right) ... dK\left(v_N\right)$$

We now impose the parametric assumptions $H(z) = 1 - (z/z_0)^{-\zeta}$ and $\tilde{\lambda} = \lambda z_0^{-\zeta}$. As shown in Appendix F, in the limiting economy as $z_0 \to 0$, the fraction of tasks with techniques delivering efficiency no greater than v converges to a Frechet distribution

$$\begin{aligned} K(v) &= e^{-\theta v^{-\zeta}} \\ \theta &= \lambda \theta^{\alpha} \Gamma\left(1 - \frac{\alpha}{N}\right)^{N} \end{aligned}$$

The expression for F is messier, but we can derive an expression for aggregate productivity $Q = \left(\int_0^{\infty} q^{\varepsilon-1} dF(q)\right)^{\frac{1}{\varepsilon-1}}$:

$$Q = \Gamma\left(1 - \frac{(\varepsilon-1)/\zeta}{N}\right)^{\frac{N}{\varepsilon-1}} \lambda^{\frac{1}{(1-\alpha)\zeta}} \Gamma\left(1 - \frac{\alpha}{N}\right)^{N \frac{1}{(1-\alpha)\zeta}}$$

This generalizes the formula for aggregate productivity with a single input given by equation (12).Since $\Gamma\left(1 - \frac{\alpha}{N}\right)^{N}$ and $\Gamma\left(1 - \frac{(\varepsilon-1)/\zeta}{N}\right)^{\frac{N}{\varepsilon-1}}$ are decreasing in N, aggregate productivity is declining in the number of inputs, as shown in Figure 9.

Aggregate productivity declines with the number of inputs as a consequence because the tail of superstar suppliers shrinks. This happens because of two mutually reinforcing mechanisms. First, the distribution of efficiency $F(q)$ becomes more concentrated. With a single input, a firm would end up with a high efficiency with one great technique. With more inputs, a firm that gets lucky enough to draw a great technique for one of its tasks may only have low efficiency techniques for its

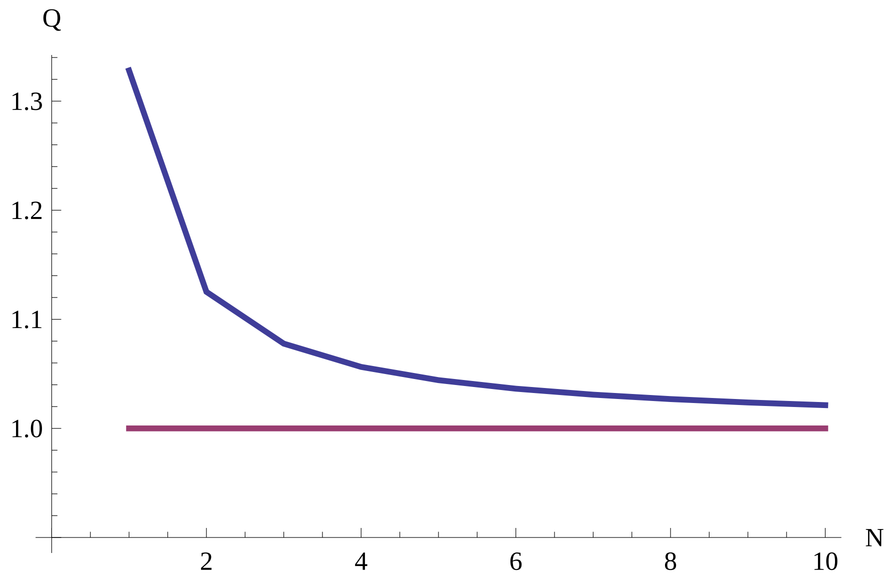


Figure 9: Aggregate Productivity by Number of Inputs

Figure 9 shows aggregate productivity Q the as the number of inputs changes. The parameters are $\alpha = 1/2$, $\zeta = 4$, and $\epsilon = 4$, with λ set so that the limiting aggregate productivity when the number of inputs grows large is 1. The purple line is the asymptotic aggregate productivity as the number of inputs grows large.

other tasks. The more inputs there are, the harder it is for a firm to have a high overall efficiency, as the the law of large numbers (across tasks) begins to kick in.

On top of this, when efficiencies are more concentrated among potential suppliers, it becomes harder to draw a high efficiency technique for a single task. The efficiency provided by a technique is zq^{α}, and less dispersion in q implies that relatively high efficiency techniques are drawn less frequently.

A stark feature of the model that drives this result is that all tasks enter symmetrically into production. If instead production were dominated by relatively few core tasks, then the distribution of efficiencies would be less concentrated and superstar suppliers would be more relevant.

Alternatively, if tasks are substitutes then the high efficiency tasks will have a large cost share. In the limit as the elasticity of substitution goes to infinity, the N input model would converge to the single input model.

7 Conclusion

This paper has described a tractable model of the formation and evolution of chains of production. The model aggregates easily, with a simple formula connecting features of the network to aggregate output. A key feature driving aggregate productivity is whether the low cost producers are able to become the important suppliers of the economy. To become an important supplier, a firm must have many potential customers who know how to use the firm's good and to be able to charge those potential customers a low price. When intermediate goods are more important in production (relative to labor), the ability to charge a low price becomes more important in winning customers, and the lowest cost firms are more likely to become superstar suppliers. This increases aggregate productivity and also increases the market concentration in sales of intermediate goods.

With more interesting network configurations, the model can be solved almost as easily. With multiple types, aggregate productivity depends on how productivity and demand covary across groups. Holding the total quantity of techniques fixed, concentration of link formation among a subset of firms, e.g., in a city, leads to positive correlation and higher aggregate output. In contrast, negative correlation, e.g., if some firms specialize in R&D whereas others specialize in marketing, can lead to significantly lower output.

Appendix

A Notation

A chain of techniques is a sequence of techniques (finite of infinite) $\phi_0\phi_1\phi_2...$ with the $s(\phi_{k+1}) = b(\phi_k)$. Given the set of all existing techniques, Φ, we can define several objects.

Let $C(j)$ be the set of chains of techniques with the property $\phi_0\phi_1\phi_2...$ with the additional property that $b(\phi_0) = j$. These are the distinct supply chains available that could be used to produce good j. For example, if firm j_1 has access to a single technique ϕ_1 with supplier j_2, and j_2 has access to two techniques ϕ_2 and ϕ_3, then C contains the three distinct chains ϕ_1, $\phi_1\phi_2$, and $\phi_1\phi_3$.

Let $C_N(j) \subseteq C(j)$ be the set of chains of techniques to produce good j with exactly N techniques. Let $b_N(j) \equiv |C_N(j)|$ be the number of such distinct supply chains of length N.

Lastly, let $C_\infty(j) \subseteq C(j)$ be the set of infinite chains of techniques to produce good j. These are all of the viable supply chains to produce good j. $C_N^\infty(j)$ is the set of chains of techniques of length N that form the beginning of an infinite supply chains. For example, $\phi_0\phi_1\phi_2... \in C^\infty$(j), then $\phi_0\phi_1$ is in $C_2^\infty(j)$.

B The Number of Supply Chains

For completeness, we show the derivation of two results from the theory of branching processes that will be used in this paper (see for example Athreya and Ney (1972)).

Let $p(k)$ be the probability that a firm has exactly k techniques, in this case equal to $\frac{e^{-\tilde{\lambda}}\tilde{\lambda}^k}{k!}$, and $P_N(l,k)$ be the probability that, in total, l different firms have among them k supply chains of length N. Note that $P_N(1,k)$ is the probability a firm has exactly k supply chains of length N (i.e., the probability that $b_N = k$).

Define $\varphi(x) = \sum_{k=0}^{\infty} p(k)x^k$ to be the probability generating function for the random variable b_1. In this case $\varphi(x) = e^{-\tilde{\lambda}(1-x)}$. Also, for each N, let $\varphi_N(\cdot)$ be the probability generating function associated with b_N. If $\varphi^{(N)}$ is the N-fold composition of φ then we have the convenient result:

Claim 1 *$\varphi_N(x) = \varphi^{(N)}(x)$*

Proof. We proceed by induction. By definition, the statement is true for $N = 1$. Noting that $\sum_{k=0}^{\infty} P_1(l,k)x^k = \varphi(x)^l$, we have

$$\begin{aligned}
\varphi_{N+1}(x) &= \sum_{l=0}^{\infty} P_{N+1}(1,l)x^l = \sum_{l=0}^{\infty}\sum_{k=0}^{\infty} P_N(1,k)P_1(k,l)x^l = \sum_{k=0}^{\infty} P_N(1,k)\sum_{l=0}^{\infty} P_1(k,l)x^l \\
&= \sum_{k=0}^{\infty} P_N(1,k)\varphi(x)^k = \varphi_N(\varphi(x))
\end{aligned}$$

■

We immediately have the following:

Claim 2 *For any x, $\mathbb{E}\left[x^{b_N}\right] = \varphi^{(N)}(x)$.*

Proof. $\mathbb{E}\left[x^{b_N}\right] = \sum_{k=0}^{\infty} P_N(1,k)x^k = \varphi_N(x) = \varphi^{(N)}(x)$ ■

We next study the probability that a single firm has no chains that continue indefinitely. This can be characterized as follows:

Claim 3 *The probability that a single firm has no chains that continue indefinitely is the smallest root, ρ, of $y = \varphi(y)$.*

Proof. The probability that a firm has no chains greater than length N is given by $P_N(1,0)$, or equivalently $\varphi^{(N)}(0)$. Then the probability that a single firm has no chains that continue indefinitely is $\lim_{N\to\infty} \varphi_N(0)$.

Next note that that φ is increasing and convex, $\varphi(1) = 1$, and $\varphi(0) \geq 0$. This implies that in the range $[0,1]$, the equation $\varphi(y) = y$ has a either a unique root at $y = 1$ or two roots, $y = 1$ and a second in $(0,1)$.

Let ρ be the smallest root. Noting that for $y \in [0,\rho)$, $y < \varphi(y) < \rho$, while for $y \in (\rho,1)$ (if such y exist), $\rho < \varphi(y) < y$. Together these imply that if $y \in [0,1)$, the sequence $\{\varphi_N(y)\}$ is monotone and bounded, and therefore has a limit. We have $\varphi_{N+1}(0) = \varphi(\varphi_N(0))$. Taking limits of both sides (and noting that φ is continuous) gives $\lim_{N\to\infty} \varphi_{N+1}(0) = \varphi(\lim_{N\to\infty} \varphi_N(0))$. Therefore the limit must be a root of $y = \varphi(y)$, and therefore must be ρ. In other words, it must be that $\lim_{N\to\infty} \varphi_N(0) = \rho$. ■

Claim 4 *If $\tilde{\lambda} \leq 1$ then with probability 1 the firm has no chains that continue indefinitely. If $\tilde{\lambda} > 1$ then there is a strictly positive probability the firm has a chain that continues indefinitely.*

Proof. In this case, we have $\varphi(x) = e^{-\tilde{\lambda}(1-x)}$. If $\tilde{\lambda} \leq 1$ then the smallest root of $y = \varphi(y)$ is $y = 1$. If $\tilde{\lambda} > 1$ the smallest root is strictly less than 1. ■

C Existence and Characterization of the Solution to the Planner's Problem

The strategy is as follows: We define a sequence of random variables $\{X_N\}_{N\in\mathbb{N}}$ with the property that the maximum feasible efficiency of a firm is given by the limit of this sequence, if such a limit exists. We then show that X_N converges to a random variable X^{sp} in $L^{\varepsilon-1}$. Next we show that the distribution of X^{sp} is given by the unique fixed point of T in $\mathcal{F}$, a subset of $\bar{\mathcal{F}}$ (and of course that such a fixed point exists). Letting F^{sp} be this least fixed point, we apply the law of large numbers for continuum of random variables (Uhlig (1996)) to argue that the cross-sectional distribution of efficiencies (in $L^{\varepsilon-1}$) is given by F^{sp} and that aggregate productivity is simply $\|X^{sp}\|_{\varepsilon-1}$.

C.1 A Unique Fixed Point

We begin by defining three functions, $\bar{f}$, f^1, and $\underline{f}$, in $\bar{\mathcal{F}}$. To do so, we define several objects that will parameterize these functions. Let $\rho \in (0,1)$ be the smallest root of $\rho = e^{-\tilde{\lambda}(1-\rho)}$. In the definition of $\bar{f}$, q_2 and β are defined as follows: Let $\beta > \varepsilon - 1$, let q_2 be defined so that $q_2^{1-\alpha} > \max\left\{1, \tilde{\lambda}^{\frac{\alpha}{\beta}} \exp \int_{z_0}^{\infty} \log(z) dH(z)\right\}$. In the definition of $\underline{f}$, $q_0 = z_0^{\frac{1}{1-\alpha}}$. The three functions are:

$$\bar{f}(q) \equiv \begin{cases} \rho, & q < q_2 \\ 1-(1-\rho)\left(\frac{q}{q_2}\right)^{\beta}, & q \geq q_2 \end{cases}$$
$$f^1(q) \equiv \begin{cases} \rho, & q < 1 \\ 1, & q \geq 1 \end{cases}$$
$$\underline{f}(q) \equiv \begin{cases} \rho, & q < q_0 \\ 1, & q \geq q_0 \end{cases}$$

Consider the partial order on $\bar{\mathcal{F}}$, the set of right continuous, weakly increasing functions $f : \mathbb{R}^+ \to [0,1]$ given by the binary relation $\preceq$: $f_1 \preceq f_2 \Leftrightarrow f_1(q) \leq f_2(q) \forall q \geq 0$. Clearly $\bar{f} \preceq f^1 \preceq \underline{f}$. Let $\mathcal{F} \subset \bar{\mathcal{F}}$ be the subset of set of nondecreasing functions $f : \mathbb{R}^+ \to [0,1]$ that satisfy $\bar{f} \preceq f \preceq \underline{f}$.

Lemma 1 *$T\underline{f} \preceq \underline{f}$ and $\bar{f} \preceq T\bar{f}$*

Proof. We first show that $T\underline{f} \preceq \underline{f}$. For $q \geq q_0$ this is immediate, as $T\underline{f}(q) \leq 1 = \underline{f}(q)$. For $q < q_0$, we have

$$T\underline{f}(q) = e^{-\tilde{\lambda}\int_0^\infty \left[1-\underline{f}\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right)\right]dH(z)} = e^{-\tilde{\lambda}\int_{q/q_0^\alpha}^\infty [1-\rho]dH(z)} \leq e^{-\tilde{\lambda}[1-\rho]\left(1-H\left(q_0^{1-\alpha}\right)\right)} = \rho = \underline{f}(q)$$

We now proceed to $\bar{f}$. First, for $q < q_2$, we have $T\bar{f}(q) = e^{-\tilde{\lambda}\int_0^\infty (1-\bar{f})dH(z)} \geq e^{-\tilde{\lambda}(1-\rho)} = \rho = f(q)$. Next, as an intermediate step, we will show that for $q \geq q_2$:

$$\left(\frac{q}{q_2}\right)^{-\beta} > \tilde{\lambda}\left[\exp\int_0^{q/q_2^\alpha} \ln\left(z^{\frac{\beta}{\alpha}}\right) dH(z) - \ln\left(\left(q/q_2^\alpha\right)^{\frac{\beta}{\alpha}}\right) H\left(q/q_2^\alpha\right)\right]$$

We start with $q_2^{1-\alpha} > \tilde{\lambda}^{\frac{\alpha}{\beta}} \exp\int_0^\infty \ln(z)\, dH(z)$ which can be written as

$$q_2^{1-\alpha} > \tilde{\lambda}^{\frac{\alpha}{\beta}} \exp\left[\int_0^{q/q_2^\alpha} \ln(z)\, dH(z) + \int_{q/q_2^\alpha}^\infty \ln(z)\, dH(z)\right]$$

Using the fact that $q \geq q_2$ and that $\int_x^\infty \ln(z) dH(z) \geq \ln(x)[1-H(x)]$ gives

$$q^{1-\alpha} > \tilde{\lambda}^{\frac{\alpha}{\beta}}\left[\exp\int_0^{q/q_2^\alpha} \ln(z)\, dH(z) + \ln\left(q/q_2^\alpha\right)\left[1-H\left(q/q_2^\alpha\right)\right]\right]$$

Raising both sides to $\frac{\beta}{\alpha}$ and dividing $\left(q/q_2^\alpha\right)^{\frac{\beta}{\alpha}}$ gives the desired result.

Next, beginning with $1 - T\bar{f}(q) \leq -\ln T\bar{f}(q)$

$$\begin{aligned}
\frac{1-T\bar{f}\left(q\right)}{1-\rho} &\leq \tilde{\lambda}\int_0^\infty \frac{1-\bar{f}\left(\left(q/z\right)^{\frac{1}{\alpha}}\right)}{1-\rho}dH\left(z\right) \\
&\leq \tilde{\lambda}\exp\int_0^\infty \ln\left[\frac{1-\bar{f}\left(\left(q/z\right)^{\frac{1}{\alpha}}\right)}{1-\rho}\right]dH\left(z\right) \\
&= \tilde{\lambda}\exp\int_0^{q/q_2^\alpha} \ln\left(\frac{\left(q/z\right)^{\frac{1}{\alpha}}}{q_2}\right)^{-\beta}dH\left(z\right) \\
&= \tilde{\lambda}\exp\left[\int_0^{q/q_2^\alpha}\ln\left(z^{\frac{\beta}{\alpha}}\right)dH\left(z\right) - \ln\left(\left(q/q_2^\alpha\right)^{\frac{\beta}{\alpha}}\right)H\left(q/q_2^\alpha\right)\right] \\
&< \left(\frac{q}{q_2}\right)^{-\beta} \\
&= \frac{1-\bar{f}\left(q\right)}{1-\rho}
\end{aligned}$$

This then gives, for $q \geq q_2$, $T\bar{f}(q) \geq \bar{f}(q)$. ■

Lemma 2 *There exist least and greatest fixed points of the operator T in $\mathcal{F}$, given by $\lim_{N\to\infty} T^N\bar{f}$ and $\lim_{N\to\infty} T^N\underline{f}$ respectively.*

Proof. Proof: The operator T is order preserving, and the $\mathcal{F}$ is a complete lattice. By the Tarski fixed point theorem, the set of fixed pints of T in $\mathcal{F}$ is also a complete lattice, and hence has a least and a greatest fixed point given by $\lim T^N\bar{f}$ and $\lim T^N\underline{f}$ respectively. ■

C.2 Existence of a Limit

We first show that the planer's problem is well defined. Given the set of techniques at a point in time, Φ, let C_j^∞ be the set of distinct infinite chains available to produce good j. For each chain $c \in C_j^\infty$ and each $n \geq 0$, let $z_n(c)$ is the productivity of the nth technique in chain c. In other words, for a chain c of techniques $\phi_0\phi_1\phi_2...$ with ϕ_0 furthest downstream, $z_n(c) = z(\phi_n)$. Finally let $\mathfrak{q}(c) \equiv \prod_{n=0}^\infty z_n(c)^{\alpha^n}$.

We want to define q_j to be the efficiency provided by the most cost effective supply chain available to produce j, or more formally that $q_j = \sup_{c\in C_j^\infty}\mathfrak{q}(c)$. To do this, we first argue that for each c, $\mathfrak{q}(c)$ is well defined. The concern is that for some c the sequence $\left\{\prod_{n=0}^N z_n(c)^{\alpha^n}\right\}$ might not converge. The following proposition shows that it does.

Define $\mathfrak{q}_N(c) \equiv \prod_{n=0}^{N-1} z_n(c)^{\alpha^n}$ for $N \geq 1$ with $\mathfrak{q}_0(c) \equiv 1$.

Lemma 3 *Assume $z_0 > 0$. Then for each $c \in C_j^\infty$, $\lim_{N\to\infty}\mathfrak{q}_N(c)$ exists.*

Proof. For each n, we can decompose $\log z_n(c)$ into $\log z_n^+(c) - \log z_n^-(c)$, where $\log z_n^+(c) = \max\{\log z_n(c), 0\}$ and $\log z_n^-(c) = \max\{-log z_n(c), 0\}$. We then have $\log \mathfrak{q}_N(c) = \sum_{n=0}^{N-1}\alpha^n \log z_n^+(c) -$

$\sum_{n=0}^{N-1} \alpha^n \log z_n^-(c)$. $\sum_{n=0}^{N-1} \alpha^n \log z_n^+(c)$ is a monotone sequence so it converges to a (possibly infinite) limit. $\sum_{n=0}^{N-1} \alpha^n \log z_n^-(c)$ is a monotone sequence bounded by $\frac{\log(1/z_0)}{1-\alpha}$ so it converges to a limit in the range $\left[0, \frac{\log(1/z_0)}{1-\alpha}\right]$. Therefore $\mathfrak{q}_N(c)$ converges to a (possibly infinite) limit. ■

Next, we define the random variable $X_{N,j} = \max_{c \in C_j^\infty} \mathfrak{q}_N(c)$. Roughly, the remainder of this subsection shows that $X_{N,j}$ converges q_j. Since $q_j = \sup_{c \in C_j^\infty} \lim_{N\to\infty} \mathfrak{q}_N(c)$, we are essentially proving that the limit can be passed through the sup.

A useful property of the variable $X_{N,j}$ is that its CDF is $T^N f^1$. It will also be useful to define the random variables $\bar{Y}_{N,j}$ and $\underline{Y}_{N,j}$. These variables will be constructed so that their CDFs are $T^N \bar{f}$, and $T^N \underline{f}$ respectively.

To do this, we expand the probability space as follows: For each realization of Φ, let $\tilde{q}(c)$ be a random variable with CDF $\frac{\bar{f}-\rho}{1-\rho}$ drawn independently for each $c \in C_{N,j}^\infty$.

For $N \geq 1$, let $\bar{Y}_{N,j} \equiv \max_{c \in C_{N,j}^\infty} \mathfrak{q}_N(c)\tilde{q}(c)^{\alpha^N}$ and $\underline{Y}_{N,j} \equiv \max_{c \in C_{N,j}^\infty} \mathfrak{q}_N(c) q_0^{\alpha^N}$. We also define $\bar{Y}_{0,j}$ and $\underline{Y}_{0,j}$ to have CDFs $\bar{f}$ and $\underline{f}$ respectively.

To improve readability, the subscript j will be suppressed when not necessary.

Lemma 4 *$\{X_N\}_{N\in\mathbb{N}}$, $\{\bar{Y}_N\}_{N\in\mathbb{N}}$, and $\{\underline{Y}_N\}_{N\in\mathbb{N}}$ are uniformly integrable in $L^{\varepsilon-1}$.*

Proof. First, Recall that $\bar{Y}_0$ is defined to have the cumulative distribution $\bar{f}$. Since the T is order preserving, the relations $T^N f^1 \succeq T^N \bar{f}$ and $T^N \bar{f} \succeq T^{N-1} \bar{f}$ imply that $T^N f^1 \succeq \bar{f}$. As a consequence, $\bar{Y}_0$ first-order stochastically dominates each X_N and $\bar{Y}_N$, and, by the identical argument, $\underline{Y}_N$. Therefore $\mathbb{E}\left|\bar{Y}_0\right|^{\varepsilon-1} = \frac{q_2^{\varepsilon-1}}{1-\frac{\varepsilon-1}{\beta}} < \infty$ serves as a uniform bound on each $\mathbb{E}\left|X_N\right|^{\varepsilon-1}$, $\mathbb{E}\left|\bar{Y}_N\right|^{\varepsilon-1}$, and $\mathbb{E}\left|\underline{Y}_N\right|^{\varepsilon-1}$. ■

Lemma 5 *There exists a random variable X^{sp} such that X_N converges to X^{sp} almost surely and in $L^{\varepsilon-1}$.*

Proof. Let $M_N \equiv \frac{X_N}{\prod_{n=0}^N \mu_n}$ where $\mu_n \equiv \int_{z_0}^\infty w^{\alpha^n} \frac{\tilde{\lambda}(1-\rho)H'(w)e^{-\tilde{\lambda}(1-\rho)[1-H(w)]}}{1-e^{-\tilde{\lambda}(1-\rho)}} dw$. We first show that M_N is a submartingale that is uniformly integrable in $L^{\varepsilon-1}$.

To do this, we define a set $\mathcal{D}_N$ as follows: Let $c_N^* = \arg\max_{c \in C_N^\infty} \mathfrak{q}_N(c)$ so that $X_N = \mathfrak{q}_N(c_N^*)$. Let $\mathcal{D}_N \subseteq C_N^\infty$ be the set of chains in C_N^∞ for which the first $N-1$ links are c_{N-1}^*. In other words, all chains in $\mathcal{D}_N$ are of the form $c_{N-1}^*\phi$ for some ϕ.

Define the random variable $D_N = \max_{c \in \mathcal{D}_N} \mathfrak{q}_N(c)$. Since $\mathcal{D}_N \subseteq C_N^\infty$, it must be that $X_N \geq D_N$. We now show that $E\left[D_N | C_{N-1}\right] \geq \mu_{N-1} X_N$:

The number of chains in $\mathcal{D}_N$ is at least one. The probability that $|\mathcal{D}_N| = k$ is $\frac{e^{-\tilde{\lambda}[1-\rho]}\left[\tilde{\lambda}(1-\rho)\right]^k}{\left[1-e^{-\tilde{\lambda}[1-\rho]}\right]k!}$ for $k \geq 1$. To see this, note that for any node, the number of techniques is poisson with mean $\tilde{\lambda}$. Each of those has probability $1-\rho$ of being viable (having a chain that continues infinitely), and we are conditioning on at least one viable technique.

Each of those techniques has a productivity drawn from H. For any ϕ such that $c_{N-1}^*\phi \in \mathcal{D}_N$, we have that

$$\Pr\left(\mathfrak{q}_N\left(c_N^*\phi\right) < x | C_{N-1}^\infty\right) = \Pr\left(z(\phi)^{\alpha^n} < x/X_{N-1}\right) = H\left(\left(x/X_{N-1}\right)^{\alpha^{-N}}\right)$$

Therefore the probability that all k of the chains in $\mathcal{D}_N$ give $\mathfrak{q}_N(c_N^*\phi) < x$ given X_{N-1} is

$$\Pr\left(D_N < x | C_{N-1}^{\infty}, |\mathcal{D}_N| = k\right) = H\left((x/X_{N-1})^{\alpha^{-N}}\right)^k$$

With this, the CDF of D_N, given $C_{N_1}^{\infty}$, is

$$\begin{aligned}
\Pr\left(D_N < x | C_{N-1}^{\infty}\right) &= \sum_{k=1}^{\infty} \Pr\left(D_N < x | X_{N-1}, |\mathcal{D}_N| = k\right) \Pr\left(|\mathcal{D}_N| = k\right) \\
&= \sum_{k=1}^{\infty} H\left((x/X_{N-1})^{\alpha^{-N}}\right)^k \frac{e^{-\tilde{\lambda}[1-\rho]}\left[\tilde{\lambda}(1-\rho)\right]^k}{\left[1-e^{-\tilde{\lambda}[1-\rho]}\right]k!} \\
&= \frac{e^{\tilde{\lambda}[1-\rho]H\left((x/X_{N-1})^{\alpha^{-N}}\right)} - 1}{e^{\tilde{\lambda}[1-\rho]} - 1}
\end{aligned}$$

We can now compute the conditional expectation of D_N (using the change of variables $w = (x/X_{N-1})^{\alpha^{-N}}$):

$$\begin{aligned}
\mathbb{E}\left[D_N | C_{N-1}^{\infty}\right] &= \int_0^{\infty} x\alpha^{-N}\frac{1}{X_{N-1}}\left(\frac{x}{X_{N-1}}\right)^{\alpha^{-N}-1}\tilde{\lambda}(1-\rho)H'\left((x/X_{N-1})^{\alpha^{-N}}\right)\frac{e^{\tilde{\lambda}[1-\rho]H\left((x/X_{N-1})^{\alpha^{-N}}\right)}}{e^{\tilde{\lambda}[1-\rho]}-1}dx \\
&= X_{N-1}\int_0^{\infty} w^{\alpha^N}\tilde{\lambda}(1-\rho)H'(w)\frac{e^{\tilde{\lambda}[1-\rho]H(w)}}{e^{\tilde{\lambda}[1-\rho]}-1}dw \\
&= \mu_N X_{N-1}
\end{aligned}$$

Putting this together, we have

$$\mathbb{E}\left[M_N | C_{N-1}^{\infty}\right] = \frac{1}{\prod_{n=0}^N \mu_n}\mathbb{E}\left[X_N | C_{N-1}^{\infty}\right] \geq \frac{1}{\prod_{n=0}^N \mu_n}\mathbb{E}\left[D_N | C_{N-1}^{\infty}\right] = \frac{1}{\prod_{n=0}^N \mu_n}\mu_N X_{N-1} = M_{N-1}$$

We next need to show that $\{M_N\}$ is uniformly integrable, i.e., that $\sup_N \mathbb{E}[M_N] < \infty$. We know that $\sup_N \mathbb{E}[X_N] < \infty$. So it suffices to show a uniform lower bound on $\prod_{n=0}^N \mu_n$. Since each $\mu_n \geq z_0^{\alpha^n}$ and $z_0 < 1$, we have that $\prod_{n=0}^N \mu_n \geq \prod_{n=0}^N z_0^{\alpha^n} \geq \prod_{n=0}^{\infty} z_0^{\alpha^n} = z_0^{\frac{1}{1-\alpha}}$.

We have therefore established that $\{M_N\}_{N\in\mathbb{N}}$ is a uniformly integrable (in L^1) submartingale, so by the martingale convergence theorem, there exists an M such that M_N converges to M almost surely. By the continuous mapping theorem, there exists an X^{sp} such that X_N converges to X^{sp} almost surely. Since each $X_N^{\varepsilon-1}$ is dominated by the integrable random variable $\bar{Y}_0^{\varepsilon-1}$, by dominated convergence we have that X_N converges to X^{sp} in $L^{\varepsilon-1}$. ■

Proposition 5 *With probability one,* $X^{sp} = \sup_{c\in C^{\infty}} \mathfrak{q}(c)$

Proof. We first show that $X^{sp} \geq \sup_{c\in C^{\infty}} \mathfrak{q}(c)$ with probability one. Consider any realization of techniques, Φ. For any $\nu > 0$, there exists a $c^* \in C^{\infty}$ such that $\mathfrak{q}(c^*) > \sup_{c\in C^{\infty}} \mathfrak{q}(c) - \nu$. There also exists an N_1 such that $N > N_1$ implies $\mathfrak{q}_N(c^*) > \mathfrak{q}(c^*) - \nu$. Lastly, with probability one there exists an N_2 such that $N > N_2$ implies $X_N < X^{sp} + \nu$. We then have for $N > \max N_1, N_2$ that

$$X^{sp} > X_N - \nu = \max_{c \in C_N^\infty} \mathfrak{q}_N(c) - \nu \geq \mathfrak{q}_N(c^*) - \nu > \mathfrak{q}(c^*) - 2\nu > \sup_{c \in C^\infty} \mathfrak{q}(c) - 3\nu, \qquad w.p.1$$

This is true for any $\nu > 0$, so $X^{sp} \geq \sup_{c \in C^\infty} \mathfrak{q}(c)$. We next show the opposite inequality. For any N, we have

$$\sup_{c \in C^\infty} \mathfrak{q}(c) \geq \sup_{c \in C^\infty} \mathfrak{q}_N(c) z_0^{\frac{\alpha^N}{1-\alpha}} = X_N z_0^{\frac{\alpha^N}{1-\alpha}}$$

Since this is true for any N and $\lim_{N \to \infty} z_0^{\frac{\alpha^N}{1-\alpha}} = 1$, we can take the limit to get $\sup_{c \in C^\infty} \mathfrak{q}(c) \geq X^{sp}$ with probability one. ■

C.3 Characterization of the Limit

We will show below that $\log \bar{Y}_N - \log \underline{Y}_N$ converges to 0 in probability. Since $X_N \in [\underline{Y}_N, \bar{Y}_N]$, it must be that both $\bar{Y}_N$ and $\underline{Y}_N$ converge to X^{sp} in probability. Convergence in probability implies convergence in distribution, which gives to implications. First, $T^N \bar{f}$ and $T^N \underline{f}$ converge to the same limiting function. Since these are the least and greatest fixed points of T in $\mathcal{F}$, this limiting function, f^{sp}, is the unique fixed point of T in $\mathcal{F}$. Second, since $T^N \bar{f} \preceq T^N f^1 \preceq T^N \underline{f}$, f^{sp} is the CDF of X^{sp}.

Lemma 6 *For any $\eta > 1$, $\lim_{N \to \infty} 1 - \varphi^{(N)} \left(\frac{\bar{f}\left(\eta^{\alpha^{-N}} q_0\right) - \rho}{1-\rho} \right) = 0$*

Proof. We first show that for $x \in [0,1]$, $\frac{d}{dx}\varphi^{(N)}(x) \leq \tilde{\lambda}^N$. To see this, note that φ is convex and $\varphi'(1) = \tilde{\lambda}$, so that $\varphi'(x) \leq \tilde{\lambda}$ for $x \leq 1$. In addition, if $x \in [0,1]$ then $\varphi(x) \in (0,1]$, which implies $\varphi^{(N)}(x) \in (0,1]$ for each N. We then have

$$\frac{d}{dx}\varphi^{(N)}(x) = \prod_{n=1}^{N} \varphi' \left(\varphi^{(n-1)}(x) \right) \leq \tilde{\lambda}^N$$

With this, for any x, we can bound $\varphi^{(N)}(x)$ by

$$\begin{aligned} \varphi^{(N)}(x) &= \varphi^{(N)}(1) - \int_x^1 \varphi^{(N)\prime}(w) dw \\ &\geq 1 - \tilde{\lambda}^N \int_x^1 dw \\ &\geq 1 - \tilde{\lambda}^N [1 - x] \end{aligned}$$

To complete the proof, we show that $\lim_{N\to\infty}\tilde{\lambda}^N\left[1-\frac{\bar{f}\left(\eta^{\alpha^{-N}}q_0\right)-\rho}{1-\rho}\right]=0$. Recall that $\frac{\bar{f}(q)-\rho}{1-rho}=\begin{cases}0 & q<q_2\\ 1-\left(\frac{q}{q_2}\right)^{-\beta} & q\geq q_2\end{cases}$. We have for $q\geq q_2$

$$\lim_{N\to\infty}\tilde{\lambda}^N\left[1-\frac{\bar{f}\left(\eta^{\alpha^{-N}}q_0\right)-\rho}{1-\rho}\right]=\lim_{N\to\infty}\tilde{\lambda}^N q_2^\beta\left(\eta^{-\beta\alpha^{-N}}\right)=0$$

■

Next we use this to show that $\log\bar{Y}_N-\log\underline{Y}_N$ converges to zero in probability.

Lemma 7 *For any $\eta>1$, $\lim_{N\to\infty}\Pr\left(\bar{Y}_N/\underline{Y}_N>\eta\right)=0$.*

Proof. For any chain $c\in C_N$, define the following two objects: $\bar{\mathfrak{q}}_N(c)\equiv\mathfrak{q}_N(c)\tilde{q}(c)^{\alpha^N}$ (recall that for each $c\in C_N^\infty$ $\tilde{q}(c)$ was a random variable with CDF $\frac{\bar{f}-\rho}{1-\rho}$) and $\underline{\mathfrak{q}}_N(c)\equiv\mathfrak{q}_N(c)q_0^{\alpha^N}$ (q_0 is the same as a random variable with CDF $\frac{\underline{f}-\rho}{1-\rho}$). With these definitions, we have $\bar{Y}_{N,j}\equiv\max_{c\in C_{N,j}^\infty}\bar{\mathfrak{q}}_N(c)$ and $\underline{Y}_{N,j}\equiv\max_{c\in C_{N,j}^\infty}\underline{\mathfrak{q}}_N(c)$.

Conditional on the set of techniques Φ, we have that for each chain c in C_N^∞ (if any exist) that $\bar{\mathfrak{q}}_N(c)/\underline{\mathfrak{q}}_N(c)=\left(\frac{\tilde{q}(c)}{q_0}\right)^{\alpha^N}$. We therefore have:

$$\Pr\left(\bar{\mathfrak{q}}_N(c)/\underline{\mathfrak{q}}_N(c)<\eta|\Phi\right)=\Pr\left(\left(\frac{\tilde{q}(c)}{q_0}\right)^{\alpha^N}<\eta|\Phi\right)=\frac{\bar{f}\left(\eta^{\alpha^{-N}}q_0\right)-\rho}{1-\rho}$$

If there are b_N such chains of length N, the probability that every one of them of them gives $\bar{\mathfrak{q}}_N(c)/\underline{\mathfrak{q}}_N(c)<\eta$ is $\left(\frac{\bar{f}\left(\eta^{\alpha^{-N}}q_0\right)-\rho}{1-\rho}\right)^{b_N}$ so that

$$\Pr\left(\bar{Y}_N/\underline{Y}_N<\eta|\Phi\right)\geq\left(\frac{\bar{f}\left(\eta^{\alpha^{-N}}q_0\right)-\rho}{1-\rho}\right)^{b_N}$$

Recall from Appendix B that for any x, $\mathbb{E}\left[x^{b_N}\right]=\varphi^{(N)}(x)$ where $\varphi^{(N)}$ is the N-fold composition of φ and the expectations are taken over realizations of Φ. We can use this to show

$$\begin{aligned}\Pr\left(\bar{Y}_N/\underline{Y}_N<\eta\right) &= \mathbb{E}\left[\Pr\left(\bar{Y}_N/\underline{Y}_N<\eta|\Phi\right)\right]\\ &\geq \mathbb{E}\left[\left(\frac{\bar{f}\left(\eta^{\alpha^{-N}}q_0\right)-\rho}{1-\rho}\right)^{b_N}\right]=\varphi^{(N)}\left(\frac{\bar{f}\left(\eta^{\alpha^{-N}}q_0\right)-\rho}{1-\rho}\right)\end{aligned}$$

Put differently, $\lim_{N\to\infty} \Pr\left(\bar{Y}_N/\underline{Y}_N > \eta\right) \leq \lim_{N\to\infty} 1 - \varphi^{(N)}\left(\frac{\bar{f}\left(\eta^{\alpha^{-N}} q_0\right)-\rho}{1-\rho}\right)$. We complete the proof by applying the previous lemma. ■

We now come to the main result.

Proposition 6 *There is a unique fixed point of T on $\mathcal{F}$, F^{sp}. F^{sp} is CDF of X^{sp}. Aggregate productivity is $Q = \left(\int_0^\infty q^{\varepsilon-1} dF^{sp}(q)\right)^{\frac{1}{\varepsilon-1}}$ with probability one.*

Proof. The combination of $\log \bar{Y}_N - \log \underline{Y}_N \xrightarrow{p} 0$, $\bar{Y}_N \geq X_N \geq \underline{Y}_N$, and $X_N \xrightarrow{p} X^{sp}$ implies that $\bar{Y}_N \xrightarrow{p} X^{sp}$ and $\underline{Y}_N \xrightarrow{p} X^{sp}$.

We first show that there is a unique fixed point, which is also the CDF of X^{sp}. The CDFs of $\bar{Y}_N$ and $\underline{Y}_N$ are $T^N \bar{f}$ and $T^N \underline{f}$ respectively. The least and greatest fixed points of T in $\mathcal{F}$ are $\lim_{N\to\infty} T^N \bar{f}$ and $\lim_{N\to\infty} T^N \underline{f}$ respectively. Convergence in probability implies convergence in distribution, so the least and greatest fixed point are the same, and that the fixed point is the CDF of X^{sp}. Call this fixed point F^{sp}.

Since $\left\{\bar{Y}_N\right\}$ and $\left\{\underline{Y}_N\right\}$ are uniformly integrable in $L^{\varepsilon-1}$, we have by Vitali's convergence theorem that $\bar{Y}^N \to X^{sp}$ in $L^{\varepsilon-1}$ and $\underline{Y}^N \to X^{sp}$ in $L^{\varepsilon-1}$.

Putting all of these pieces together, we have that the CDF of q_j is F^{sp}. We next show that aggregate productivity is the $Q = \left(\int_0^\infty q^{\varepsilon-1} df^{sp}(q)\right)^{\frac{1}{\varepsilon-1}}$. For this we simply apply the law of large numbers for a continuum economy of Uhlig (1996). To do this, we must verify that the efficiencies are pairwise uncorrelated. This is trivial: consider two firms, j and i. Since the number of firms in any of j's supply chains is countable, the probability that i and j have overlapping supply chains is zero. The theorem in Uhlig (1996) also requires that the variable in question has a finite variance, and if it does, the then the L^2 integral exists. Here we are interested in the $L^{\varepsilon-1}$ norm, so we require the X^{sp} is $L^{\varepsilon-1}$ integrable. Therefore we have that $Q = \left(\int_0^\infty q^{\varepsilon-1} df^{sp}(q)\right)^{\frac{1}{\varepsilon-1}}$ with probability one. ■

D Size Distribution

D.1 Number of Customers

Proposition 7 *Let p_n be the mass of firms with n customers. Then*

$$p_n = \int_0^\infty \frac{\left(\frac{w^{-\alpha}}{\Gamma(1-\alpha)}\right)^n e^{-\left(\frac{w^{-\alpha}}{\Gamma(1-\alpha)}\right)}}{n!} e^{-w} dw$$

The counter cumulative distribution has a tail index of $1/\alpha$.

Proof.

$$
\begin{aligned}
p_n &= \int_0^\infty \frac{\left(\frac{\lambda}{\theta} q^{\alpha\zeta}\right)^n e^{-\frac{\lambda}{\theta} q^{\alpha\zeta}}}{n!} dF(q) \\
&= \int_0^\infty \frac{\left(\frac{\lambda\theta^\alpha}{\theta}\left(\theta q^{-\zeta}\right)^{-\alpha}\right)^n e^{-\left(\frac{\lambda\theta^\alpha}{\theta}\left(\theta q^{-\zeta}\right)^{-\alpha}\right)}}{n!} \zeta\theta q^{-\zeta-1} e^{-\theta q^{-\zeta}} dq \\
&= \int_0^\infty \frac{\left(\frac{w^{-\alpha}}{\Gamma(1-\alpha)}\right)^n e^{-\left(\frac{w^{-\alpha}}{\Gamma(1-\alpha)}\right)}}{n!} e^{-w} dw
\end{aligned}
$$

The second line uses the fact that $\theta = \Gamma(1-\alpha)\lambda\theta^\alpha$. We can rewrite this as

$$
p_n = \int_0^\infty \frac{u^n e^{-n}}{n!} \frac{e^{-[\Gamma(1-\alpha)u]^{-1/\alpha}}}{\Gamma(1-\alpha)^{\frac{1}{\alpha}}\alpha} u^{-\frac{1}{\alpha}-1} du
$$

Since $\frac{e^{-[\Gamma(1-\alpha)u]^{-1/\alpha}}}{\Gamma(1-\alpha)^{\frac{1}{\alpha}}\alpha}$ is slowly varying as $u \to \infty$, applying Theorem 2.1 of Willmot (1990) gives

$$
p_n \sim \frac{e^{-[\Gamma(1-\alpha)n]^{-1/\alpha}}}{\Gamma(1-\alpha)^{\frac{1}{\alpha}}\alpha} n^{-\frac{1}{\alpha}-1}, \qquad n \to \infty
$$

Therefore the counter-cumulative distribution, $\Pr(\#\text{ customers } > n)$, has a tail index of $1/\alpha$. ■

D.2 The Distribution of Employment

Let $\chi(s)$ be the characteristic function associated with the cross sectional distribution of employment. This is the central object of interest, but to get at it, we take several intermediate steps. If $L^0(q)$ is the (deterministic) quantity of labor used producing final output for a firm with efficiency q, let $\chi^0(s|q)$ be the characteristic function for labor used for final demand. We have

$$
\begin{aligned}
\chi^0(s|q) &= \int_{-\infty}^\infty e^{isl}\delta\left(l - L^0(q)\right) dl = e^{isL^0(q)} \int_{-\infty}^\infty e^{isx}\delta(x)dx \\
&= e^{isL^0(q)}
\end{aligned}
$$

where δ is the Dirac delta function.

We showed above that the quantity of labor used by a single customer can be treated as an IID random variable. Recall also the convenient fact that if firm j uses L_j units of labor, j's supplier will use αL_j units of labor to make the inputs for j. If L_j is an IID random variable, αL_j is as well. With this in mind, let $\chi^1(s)$ be the characteristic function associated with the labor required to make the inputs for a single customer. If the density of labor is $\Pr(L_j = l)$, the density of labor used for a supplier is $\frac{1}{\alpha}\Pr\left(L_j = \frac{1}{\alpha}\right)$, so that we have

$$
\begin{aligned}
\chi^1(s) &= \int_{-\infty}^{\infty} \frac{1}{\alpha} \Pr\left(L_j = \frac{l}{\alpha}\right) e^{-isl} dl = \int_{-\infty}^{\infty} \Pr\left(L_j = \frac{l}{\alpha}\right) e^{-i(\alpha s)\frac{l}{\alpha}} d\left(\frac{l}{\alpha}\right) \\
&= \chi(\alpha s)
\end{aligned}
$$

Let $\chi^{int}(s|q)$ be the characteristic function associated with the labor used for all intermediates among firms with efficiency q. Using the fact that the characteristic function of the sum of independent random variables is the product of the characteristic functions of each of the random variables, we can write $\chi^{int}(s|q)$ as

$$
\begin{aligned}
\chi^{int}(s|q) &= \sum_{n=0}^{\infty} \chi^1(s)^n \Pr(n \text{ customers}|q) = \sum_{n=0}^{\infty} \chi^1(s)^n \frac{\left(\frac{\lambda}{\theta} q^{\alpha\zeta}\right)^n e^{-\lambda \frac{q^{\alpha\zeta}}{\theta}}}{n!} = e^{-\frac{\lambda}{\theta} q^{\alpha\zeta}\left[1-\chi^1(s)\right]} \\
&= e^{-\frac{\lambda}{\theta} q^{\alpha\zeta}[1-\chi(\alpha s)]}
\end{aligned}
$$

where the second equality uses the fact that among firms with efficiency q the distribution of customers is Poisson with parameter $\frac{\lambda}{\theta} q^{\alpha\zeta}$.

We can put these together to derive an expression for $\chi(s|q)$, the characteristic function associated with the distribution of employment among firms with efficiency q:

$$
\chi(s|q) = \chi^0(s|q)\, \chi^{int}(s|q) = e^{isL^0(q)} e^{-\frac{\lambda}{\theta} q^{\alpha\zeta}[1-\chi(\alpha s)]}
$$

Lastly we can integrate across firms, which delivers a single recursive equation that defines $\chi(s)$:

$$
\chi(s) = \int_0^{\infty} \chi(s|q)\, dF(q) = \int_0^{\infty} e^{isL^0(q)} e^{-\frac{\lambda}{\theta} q^{\alpha\zeta}[1-\chi(\alpha s)]} dF(q)
$$

We now plug in the functional forms $L^0(q) = \frac{\left(\theta q^{-\zeta}\right)^{-\frac{\varepsilon-1}{\zeta}}}{\Gamma\left(1-\frac{\varepsilon-1}{\zeta}\right)} (1-\alpha) L$ and $dF(q) = \theta\zeta q^{-\zeta-1} e^{-\theta q^{-\zeta}} dq$ to give

$$
\chi(s) = \int_0^{\infty} e^{is\left(\frac{\left(\theta q^{-\zeta}\right)^{-\frac{\varepsilon-1}{\zeta}}}{\Gamma\left(1-\frac{\varepsilon-1}{\zeta}\right)}(1-\alpha)L\right)} e^{-\mu \frac{\theta^{\alpha}\left(\theta q^{-\zeta}\right)^{-\alpha}}{\theta}[1-\chi(\alpha s)]} \theta\zeta q^{-\zeta-1} e^{-\theta q^{-\zeta}} dq
$$

and using a change of variables along with the fact that $\theta = \Gamma(1-\alpha)\lambda\theta^{\alpha}$ gives

$$
\chi(s) = \int_0^{\infty} \exp\left\{ is \frac{t^{-\frac{\varepsilon-1}{\zeta}}}{\Gamma\left(1-\frac{\varepsilon-1}{\zeta}\right)} (1-\alpha) L - \frac{t^{-\alpha}}{\Gamma(1-\alpha)} [1-\chi(\alpha s)] - t \right\} dt \tag{16}
$$

E Pairwise Stable Equilibria

We are careful here to use notation that allows for the possibility that a firm uses a production chain with a cycle: at least some of the input from j is used in as an intermediate good in the supply chain used by j. In this case, if a firm decides to lower its price of final output, it sells more final output but also raises the demand for its good as an intermediate input.

In particular, given an arrangement, for each $\phi \in \Phi_j^{B*}$, let $x_{b(\phi)}\left(y_j^0\right)$ be the quantity of good j sold to firm $b(\phi)$ given y_j^0, holding constant the final output of other firms. $x_{b(\phi)}\left(y_j^0\right)$ is an nondecreasing and weakly concave function of y_j^0. The profit for firm j can be written as:

$$\pi_j = \max_{p_j^0, y_j^0, L, x} p_j^0 y_j^0 + \sum_{\phi \in \Phi_j^{B*}} \left[p(\phi) x_{b(\phi)}\left(y_j^0\right) + \tau(\phi)\right] - wL - p\left(\phi^*\right) x - \tau\left(\phi^*\right)$$

subject to

$$\begin{aligned} y_j^0 + \sum_{\phi \in \Phi_j^{B*}} x_{b(\phi)}\left(y_j^0\right) &\leq \frac{1}{\alpha^\alpha (1-\alpha)^{1-\alpha}} z\left(\phi^*\right) x^\alpha L^{1-\alpha} \\ y_j^0 &\leq D\left(p_j^0\right) \end{aligned}$$

The first order conditions with respect to inputs deliver cost minimization, so that firm j's marginal cost is $c_j \equiv \frac{1}{z(\phi^*)} p\left(\phi^*\right)^\alpha w^{1-\alpha}$. The price of final output p_j^0 then satisfies:

$$p_j^0 = \arg\max_p \left(p - c_j\right) D(p) + \sum_{\phi \in \Phi_j^{B*}} \left(p(\phi) - c_j\right) x_{b(\phi)}\left(D(p)\right)$$

Note that if good j is in the supply chain to produce good j (there is a cycle), firm j internalizes how changing final output would affect there sales of j as an intermediate.

Proposition 8 *In any pairwise stable equilibrium, for each technique ϕ^* that is used, $p(\phi) = c_{s(\phi)}$.*

Proof. Assume there is a $\phi^* = \{j, i, z\}$ such that $x\left(\phi^*\right) > 0$ and $p\left(\phi^*\right) \neq c_i$. Consider the deviation

$$\begin{aligned} \hat{p}\left(\phi^*\right) &= c_i \\ \hat{\tau}\left(\phi^*\right) &= \tau\left(\phi^*\right) + x_j\left(p\left(\phi^*\right) - c_i\right) + K \end{aligned}$$

where $K \in \left(0, \left(\hat{p}_j^0 - \hat{c}_j\right)\hat{y}_j^0 - \left(p_j^0 - \hat{c}_j\right) y_j^0 + \sum_{\phi \in \Phi_j^{B*}} \left(p(\phi) - \hat{c}_j\right)\left[x_{b(\phi)}\left(\hat{y}_j^0\right) - x_{b(\phi)}\left(y_j^0\right)\right]\right)$. Given i's cost c_i, the cost for j of the new technique is $\hat{c}_j = \frac{1}{z(\phi^*)} c_i^\alpha w^{1-\alpha}$. The change in profit for i from the deviation is

$$\hat{\pi}_i - \pi_i = \hat{\tau}\left(\phi^*\right) - \tau\left(\phi^*\right) + \left(\hat{p}\left(\phi^*\right) - c_i\right)\hat{x}_j - \left(p\left(\phi^*\right) - c_i\right) x_j = K > 0$$

The change in profit for j from the deviation is

$$\begin{aligned} \hat{\pi}_j - \pi_j &= -\left[\hat{\tau}\left(\phi^*\right) - \tau\left(\phi^*\right)\right] + \hat{p}_j^0 \hat{y}_j^0 - p_j^0 y_j^0 + \sum_{\phi \in \Phi_j^{B*}} p(\phi)\left[x_{b(\phi)}\left(\hat{y}_j^0\right) - x_{b(\phi)}\left(y_j^0\right)\right] \\ &\quad - \left\{\left[w\hat{L}_j + c_i \hat{x}_j\right] - \left[wL_j + p\left(\phi^*\right) x_j\right]\right\} \end{aligned}$$

substituting in the expressions for $\hat{\tau}$, $\hat{p}$, and $\hat{c}$, we have

$$\begin{aligned}\hat{\pi}_j - \pi_j &= \left(\hat{p}_j^0 - \hat{c}_j\right)\hat{y}_j^0 - \left(p_j^0 - \hat{c}_j\right)y_j^0 + \sum_{\phi\in\Phi_j^{B*}}\left(p(\phi) - \hat{c}_j\right)\left[x_{b(\phi)}\left(\hat{y}_j^0\right) - x_{b(\phi)}\left(y_j^0\right)\right] - K \\ &\quad + \left[\left((1-\alpha) + \frac{c_i}{p\left(\phi^*\right)}\alpha\right)\left(\frac{p\left(\phi^*\right)}{c_i}\right)^{\alpha} - 1\right]\hat{c}_j\left(y_j^0 + \sum_{\phi\in\Phi_j^{B*}} x_{b(\phi)}\left(y_j^0\right)\right)\end{aligned}$$

We know that $[(1-\alpha) + x\alpha]\,x^{-\alpha} \geq 1$ for $\alpha \in [0,1]$ (since x^t is a convex function of t we can use Jensen's inequality to get $(1-\alpha)x^{-\alpha} + \alpha x^{1-\alpha} \geq x^{(1-\alpha)(-\alpha)+(1-\alpha)\alpha} = 1$). Therefore $\left((1-\alpha) + \frac{c_i}{p(\phi^*)}\alpha\right)\left(\frac{p(\phi^*)}{c_i}\right)^{\alpha} \geq 1$, so that

$$\hat{\pi}_j - \pi_j \geq \left(\hat{p}_j^0 - \hat{c}_j\right)\hat{y}_j^0 - \left(p_j^0 - \hat{c}_j\right)y_j^0 + \sum_{\phi\in\Phi_j^{B*}}\left(p(\phi) - \hat{c}_j\right)\left[x_{b(\phi)}\left(\hat{y}_j^0\right) - x_{b(\phi)}\left(y_j^0\right)\right] - K > 0$$

Therefore the deviation is mutually beneficial. ■

Proposition 9 *In any pairwise stable equilibrium, for each technique ϕ^* that is used, $\tau\left(\phi^*\right) \geq 0$.*

Proof. If $\tau\left(\phi^*\right) < 0$, for some $\phi^* = \{j, i, z\}$ then consider the unilateral deviation $\hat{A}\left(\phi^*\right) = 0$ (the supplier i refuses the contract).

After the deviation, every firm's input price remains the same except j. j's price will change and will use a technique from an alternative supplier (the next one on his list if there is one) and hence an alternative supply chain. The quantity produced by i will change: it will no longer produce inputs for j, and if the i is in j's alternative supply chain, i may produce more for that chain. However, since price equals marginal cost, this change in quantity produced has no effect on i's profit. We therefore have

$$\hat{\pi}_i = \pi_i - \tau\left(\phi^*\right) > \pi_i$$

which confirms the unilateral deviation. ■

Proposition 10 *In any pairwise stable equilibrium, $q_j = \max_{\phi\in\Phi_j^S}\left\{z(\phi)q_{s(\phi)}^{\alpha}\right\}$.*

Proof. Assume that $\phi_j^* = \phi^1 \equiv \left\{j, i^1, z^1\right\}$ and let $c_j = z^1 q_{i^1}^{\alpha}$. Toward a contradiction, assume there exists a $\phi^2 = \left\{j, i^2, z^2\right\} \in \Phi_j^S$ and that $c_j > \hat{c}_j \equiv z^2 q_{i^2}^{\alpha}$.

Consider the pairwise deviation: $\hat{O}_j\left(\phi^2\right) = 1$, $\hat{O}_j\left(\phi^1\right) > 1$, $\hat{A}_{i^2}\left(\phi^2\right) = 1$, $\hat{p}\left(\phi^2\right) = c_{i^2}$, and $\hat{\tau}\left(\phi^2\right) = \frac{\max_p(p-\hat{c}_j)D(p) - \max_p(p-c_j)D(p)}{2} > 0$.

In the deviation, j's marginal cost falls, but every other firm's marginal cost is unchanged. Therefore j lowers its price of final output, but no other firm does, so the quantity of final output of j rises.

Firm j finds the deviation profitable, $\hat{\pi}_j > \pi_j$, because marginal cost has fallen: (i) j now makes a positive profit on each unit of output sold as intermediate output (if j is in a cycle, then profit from intermediates would rise even more); (ii) The change in profit from final sales is greater than the fee $\hat{\tau}\left(\phi^2\right)$; and (iii) j no longer as to pay $\tau\left(\phi^1\right)$ (which is nonnegative).

Firm i^2 also finds the deviation profitable, $\hat{\pi}_{i^2} > \pi_{i^2}$, because it now collects $\hat{\tau}\left(\phi^2\right)$. Since price equals marginal cost, changes in quantity produced as intermediates have no impact on its profit.

This confirms the unilateral deviation. ■

We now show that if there if each firm has access to a technology to produce its good using only labor, $y_j = \underline{q}L$ with common productivity $\underline{q}$, then in every pairwise stable equilibrium aggregate productivity Q is the same as in the planner's problem.

If q_j^{pw} is the efficiency of firm j in a pairwise stable equilibrium, let $F^{pw} = \Pr\left(q_j^{pw} < q\right)$ be the cross sectional CDF of efficiencies in such an equilibrium. We can compare this to the distribution of efficiencies that solve the planner's problem, F^{sp}.

Proposition 11 *In any pairwise stable equilibrium, $F^{pw} = F^{sp}$ and $Q^{pw} = Q^{sp}$.*

Proof. In both the planner's problem and in any pairwise stable equilibrium, each firm chooses the technique that delivers the most cost effective combination:

$$q_j = \max\left\{\underline{q}, \max_{\phi\in\Phi_j^B}\left\{z(\phi)q_{s(\phi)}^{\alpha}\right\}\right\}$$

For each firm j, we are interested in the efficiency in the social planner's problem, q_j^{sp}, and efficiency in a pairwise equilibrium. These satisfy the following three equations:

$$\begin{aligned}
q_j^{sp} &= \max\left\{\sup_{c\in C^\infty}\mathfrak{q}(c), \sup_{N\in\mathbb{N}}\max_{c\in C_N}\mathfrak{q}_N(c)\underline{q}^{\alpha^N}\right\}, \qquad w.p.1 \\
q_j^{pw} &\geq \sup_{N\in\mathbb{N}}\max_{c\in C_N}\mathfrak{q}_N(c)\underline{q}^{\alpha^N} \\
q_j^{pw} &\leq q_j^{sp}, \qquad w.p.1
\end{aligned}$$

These can be interpreted as follows: First, the chain with the maximum feasible efficiency available to the planner is either an infinite chain or a finite chain that ends with the furthest firm upstream using the outside option. Second, in any pairwise stable equilibrium, the efficiency of a firm must be at least that of any finite chain that ends with a firm using its outside option $\underline{q}$. If otherwise, at least one of the firms in that chain is must be using a suboptimal technique. Third, efficiency in a pairwise stable equilibrium must be feasible. The first and third equations are satisfied only with probability one because the planner may deviate from the maximum feasible efficiency for a set of firms with measure zero.

The remainder of the proof follows broadly along the same lines as Appendix C with some minor modifications. As a result I will only provide a sketch, pointing out the differences. We can construct an operator in the same manner as before:

$$Tf(q) \equiv \begin{cases} 0, & q < \underline{q} \\ e^{-\tilde{\lambda}\left[1-\int_0^\infty f\left(\left(\frac{q}{z}\right)^{\frac{1}{\alpha}}\right)dH(z)\right]}, & q \geq \underline{q} \end{cases}$$

We can also define three functions.

$$\underline{f}(q) \equiv \begin{cases} 0, & q < \underline{q} \\ 1, & q \geq \underline{q} \end{cases}$$

$$f^1(q) \equiv \begin{cases} 0, & q < 1 \\ 1, & q \geq 1 \end{cases}$$

$$\bar{f}(q) \equiv \begin{cases} 0, & q < q_2 \\ 1 - \left(\frac{q}{q_2}\right)^{-\beta}, & q \geq q_2 \end{cases}$$

With the additional restriction that $q_2 > \underline{q}$.

Let $\mathcal{F}$ be the subset of $\bar{\mathcal{F}}$ defined by $\{f | \bar{f} \preceq f \preceq \underline{f}\}$. In the same manner as before, we can show that $T\bar{f} \succeq \bar{f}$ and $T\underline{f} \preceq \underline{f}$. This means that there is a least and greatest fixed point of T in $\mathcal{F}$, given by $\lim_{N\to\infty} T^N \bar{f}$ and $\lim_{N\to\infty} T^N \underline{f}$ respectively.

Define $\underline{Y}_N \equiv \max_{\hat{N} \leq N} \max_{c \in C_{\hat{N}}} \prod_{n=0}^{\hat{N}-1} \mathfrak{q}(c) \underline{q}^{\alpha^{\hat{N}}}$. With this we can define variables $X_N \equiv \max\left\{\underline{Y}_N, \max_{c\in C_N} \prod_{n=0}^{N-1} \mathfrak{q}(c)\right\}$ and $\bar{Y}_N \equiv \max\left\{\underline{Y}_N, \max_{c\in C_N} \prod_{n=0}^{N-1} \mathfrak{q}(c)\tilde{q}(c)^{\alpha^N}\right\}$, where $\tilde{q}(c)$ is a random variable drawn independently for each c in C_N for each realization of Φ. The CDFs of X_N, $\bar{Y}_N$, and $\underline{Y}_N$ are $T^N f^1$, $T^N \bar{f}$, and $T^N \underline{f}$ respectively.

Let $X^{sp} \equiv \lim_{N\to\infty} X_N$ and $X^{pw} \equiv \lim_{N\to\infty} \underline{Y}_N$ In the same manner as before we can show that these limits exists and that with probability 1 we have both $X^{sp} = q^{sp}$ and $X^{pw} \leq q^{pw} \leq X^{sp}$. Lastly, we can show that F^{sp} is the unique fixed point of T on $\mathcal{F}$, and is the CDF of both X^{sp} and X^{pw}. Applying the law of large numbers gives $Q^{sp} = Q^{pw}$. ■

F Multiple Inputs

Fraction of tasks that deliver efficiency no greater than v is $K(v)$:

$$K(v) = e^{-\tilde{\lambda} \int_0^\infty \left[1 - F\left(\left(\frac{v}{z}\right)^{1/\alpha}\right)\right] dH(z)}$$

Since $q = \prod_{n=1}^N v_n^{\frac{1}{N}}$, we the cross sectional distribution of efficiency $F(q)$ is

$$F(q) = \int_0^\infty ... \int_0^\infty K\left(\frac{q^N}{\prod_{n=2}^N v_n}\right) dK(v_2)\, dK(v_3) ... dK(v_N)$$

We now impose the parametric assumptions $H(z) = 1 - (z/z_0)^{-\zeta}$ and $\tilde{\lambda} = \lambda z_0^{-\zeta}$. This gives

$$\begin{aligned} K(v) &= e^{-\lambda \int_{z_0}^\infty \left[1 - F\left(\left(\frac{v}{z}\right)^{1/\alpha}\right)\right] \zeta z^{-\zeta-1} dz} \\ &= e^{-\lambda v^{-\zeta} \int_0^{(v/z_0)^{1/\alpha}} [1-F(r)]\alpha\zeta r^{\alpha\zeta-1} dr} \end{aligned}$$

Taking a limit as $z_0 \to 0$ gives

$$K(v) = e^{-\theta v^{-\zeta}}$$

where $\theta = \lambda \int_0^\infty [1 - F(r)]\, \alpha\zeta r^{\alpha\zeta-1} dr$

Next, we solve for θ. To do this we can rewrite $F(q)$ as (using the substitution $w_n = \theta v_n^{-\zeta}$)

$$\begin{aligned}
F(q) &= \int_0^\infty ... \int_0^\infty e^{-\theta\left(q^N/\prod_{n=2}^N v_n\right)^{-\zeta}} \theta\zeta v_2^{-\zeta-1} e^{-\theta v_2^{-\zeta}} dv_2 ... \theta\zeta v_N^{-\zeta-1} e^{-\theta v_N^{-\zeta}} dv_N \\
&= \int_0^\infty ... \int_0^\infty e^{-\left(\theta q^{-\zeta}\right)^N/\prod_{n=2}^N w_n} e^{-w_2} dw_2 ... e^{-w_N} dw_N
\end{aligned}$$

so that the density is

$$F'(q) = \frac{1}{q} N\zeta \left(\theta q^{-\zeta}\right)^N \int_0^\infty ... \int_0^\infty e^{-\left(\theta q^{-\zeta}\right)^N/\prod_{n=2}^N w_n} e^{-w_2} \frac{dw_2}{w_2} ... e^{-w_N} \frac{dw_N}{w_N}$$

We can take the expression for θ, integrate by parts, and substitute in for F' to get

$$\begin{aligned}
\theta &= \lambda \int_0^\infty \left[1 - F(r)\right] \alpha\zeta r^{\alpha\zeta-1} dr \\
&= \lambda \int_0^\infty F'(r) \, r^{\alpha\zeta} dr \\
&= \lambda \int_0^\infty \frac{1}{r} N\zeta \left(\theta r^{-\zeta}\right)^N \int_0^\infty ... \int_0^\infty e^{-\left(\theta r^{-\zeta}\right)^N/\prod_{n=2}^N w_n} e^{-w_2} \frac{dw_2}{w_2} ... e^{-w_N} \frac{dw_N}{w_N} r^{\alpha\zeta} dr
\end{aligned}$$

Making the substitution $u = \frac{\left(\theta r^{-\zeta}\right)^N}{\prod_{n=2}^N w_n}$, this becomes

$$\begin{aligned}
\theta &= \lambda\theta^\alpha \int_0^\infty \int_0^\infty ... \int_0^\infty e^{-u} e^{-w_2} \frac{dw_2}{w_2} ... e^{-w_N} \frac{dw_N}{w_N} u^{1-\frac{\alpha}{N}} \left(\prod_{n=2}^N w_n\right)^{1-\frac{\alpha}{N}} \frac{du}{u} \\
&= \lambda\theta^\alpha \Gamma\left(1 - \frac{\alpha}{N}\right)^N
\end{aligned}$$

Aggregate productivity Q can be computed as follows:

$$\begin{aligned}
Q^{\varepsilon-1} &= \int_0^\infty q^{\varepsilon-1} dF(q) \\
&= \int_0^\infty q^{\varepsilon-1} \frac{1}{q} N\zeta \left(\theta q^{-\zeta}\right)^N \int_0^\infty ... \int_0^\infty e^{-\left(\theta q^{-\zeta}\right)^N/\prod_{n=2}^N w_n} e^{-w_2} \frac{dw_2}{w_2} ... e^{-w_N} \frac{dw_N}{w_N} dq
\end{aligned}$$

Making the substitution $u = \frac{\left(\theta q^{-\zeta}\right)^N}{\prod_{n=2}^N w_n}$ and this becomes

$$\begin{aligned}
Q^{\varepsilon-1} &= \theta^{\frac{\varepsilon-1}{\zeta}} \int_0^\infty \left(\prod_{n=2}^N w_n\right)^{1-\frac{\varepsilon-1}{\zeta}/N} u^{1-\frac{\varepsilon-1}{\zeta}/N} \int_0^\infty ... \int_0^\infty e^{-u} e^{-w_2} \frac{dw_2}{w_2} ... e^{-w_N} \frac{dw_N}{w_N} \frac{du}{u} \\
&= \theta^{\frac{\varepsilon-1}{\zeta}} \Gamma\left(1 - \frac{(\varepsilon-1)/\zeta}{N}\right)^N
\end{aligned}$$

so that aggregate productivity is

$$Q = \Gamma\left(1 - \frac{(\varepsilon - 1)/\zeta}{N}\right)^{\frac{N}{\varepsilon-1}} \lambda^{\frac{1}{(1-\alpha)\zeta}} \Gamma\left(1 - \frac{\alpha}{N}\right)^{N\frac{1}{(1-\alpha)\zeta}}$$

F.1 Number of Customers

Given efficiency q_i, efficiency that a single technique that uses that good delivers efficiency less than by a technique is probability that a v is $H\left(\frac{v}{q_i^\alpha}\right)$, with density $\frac{1}{q_i^\alpha} H'\left(\frac{v}{q_i^\alpha}\right)$. The probability that the potential buyer has no other techniques better than v is simply $K(v)$ (more properly, the probability is $\frac{K(v)-e^{-\tilde{\lambda}}}{\left(1-e^{-\tilde{\lambda}}\right)\int_0^\infty F\left(\left(\frac{v}{z}\right)^{1/\alpha}\right) dH(z)}$ which converges to $K(v)$ as $z_0 \to 0$). So the probability that a technique is the potential buyer's best is $\int_0^\infty \frac{1}{q_i^\alpha} H'\left(\frac{v}{q_i^\alpha}\right) K(v) dv$. Since the number of potential buyer's follows a Poisson distribution with mean $N\tilde{\lambda}$, the number of actual buyers is Poisson with mean:

$$\begin{aligned} N\tilde{\lambda} \int_0^\infty \frac{1}{q_i^\alpha} H'\left(\frac{v}{q_i^\alpha}\right) K(v)\, dv &\to N\lambda \int_0^\infty \frac{1}{q_i^\alpha} \zeta \left(\frac{v}{q_i^\alpha}\right)^{-\zeta-1} e^{-\theta v^{-\zeta}} dv \\ &= N\frac{\lambda}{\theta} q_i^{\alpha\zeta} \end{aligned}$$

References

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "Cascades in Networks and Aggregate Volatility," NBER Working Papers 16516, National Bureau of Economic Research, Inc November 2010.

Alvarez, Fernando E., Francisco J. Buera, and Jr. Robert E. Lucas, "Models of Idea Flows," NBER Working Papers 14135, National Bureau of Economic Research, Inc June 2008.

Atalay, Enghin, Ali Hortacsu, James Roberts, and Chad Syverson, "Network Structure of Production," *Proceedings of the National Academy of Sciences*, 03 2011, *108* (13), 5199–202.

Athreya, Krishna B. and Peter Ney, *Branching processes*, Springer-Verlag, Berlin, New York,, 1972.

Barabasi, Albert-Laszlo and Reka Albert, "Emergence of Scaling in Random Networks," *Science*, 10 1999, *286*, 509–512.

Carvalho, Vasco, "Aggregate fluctuations and the network structure of intersectoral trade," Economics Working Papers 1206, Department of Economics and Business, Universitat Pompeu Fabra November 2007.

Chaney, Thomas, "The Network Structure of International Trade," NBER Working Papers 16753, National Bureau of Economic Research, Inc January 2011.

Dupor, Bill, "Aggregation and irrelevance in multi-sector models," *Journal of Monetary Economics*, April 1999, *43* (2), 391–409.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, September 2002, *70* (5), 1741–1779.

Foerster, Andrew T., Pierre-Daniel G. Sarte, and Mark W. Watson, "Sectoral vs. Aggregate Shocks: A Structural Factor Analysis of Industrial Production," NBER Working Papers 14389, National Bureau of Economic Research, Inc October 2008.

Horvath, Michael, "Cyclicality and Sectoral Linkages: Aggregate Fluctuations from Independent Sectoral Shocks," *Review of Economic Dynamics*, October 1998, *1* (4), 781–808.

Jackson, Matthew O., "A Survey of Models of Network Formation: Stability and Efficiency," Game Theory and Information 0303011, EconWPA March 2003.

_ , *Social and Economic Networks*, Princeton, NJ, USA: Princeton University Press, 2008.

_ **and Brian W. Rogers**, "Meeting Strangers and Friends of Friends: How Random Are Social Networks?," *American Economic Review*, June 2007, *97* (3), 890–915.

Jones, Charles I., "Intermediate Goods, Weak Links, and Superstars: A Theory of Economic Development," NBER Working Papers 13834, National Bureau of Economic Research, Inc March 2008.

Kelly, Morgan, "The Dynamics of Smithian Growth," *The Quarterly Journal of Economics*, August 1997, *112* (3), 939–64.

_ , "The Invention of Invention," Working Papers 200515, School Of Economics, University College Dublin September 2005.

Kortum, Samuel S., "Research, Patenting, and Technological Change," *Econometrica*, November 1997, *65* (6), 1389–1420.

Long, John B and Charles I Plosser, "Real Business Cycles," *Journal of Political Economy*, February 1983, *91* (1), 39–69.

Lucas, Robert E., "Ideas and Growth," *Economica*, 02 2009, *76* (301), 1–19.

Rosenberg, Nathan, "Technological Interdependence in the American Economy," *Technology and Culture*, 1979, *20* (1), pp. 25–50.

Uhlig, Harald, "A Law of Large Numbers for Large Economies," *Economic Theory*, 1996, *8* (1), 41–50.

Willmot, Gordon E., "Asymptotic Tail Behaviour of Poisson Mixtures with Applications," *Advances in Applied Probability*, 1990, *22* (1), pp. 147–159.

Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

Working Paper Series *(continued)*

Working Paper Series *(continued)*

The Case of the Undying Debt — **WP-09-12**
François R. Velde

Paying for Performance: The Education Impacts of a Community College Scholarship Program for Low-income Adults — **WP-09-13**
Lisa Barrow, Lashawn Richburg-Hayes, Cecilia Elena Rouse, and Thomas Brock

Establishments Dynamics, Vacancies and Unemployment: A Neoclassical Synthesis — **WP-09-14**
Marcelo Veracierto

The Price of Gasoline and the Demand for Fuel Economy: Evidence from Monthly New Vehicles Sales Data — **WP-09-15**
Thomas Klier and Joshua Linn

Estimation of a Transformation Model with Truncation, Interval Observation and Time-Varying Covariates — **WP-09-16**
Bo E. Honoré and Luojia Hu

Self-Enforcing Trade Agreements: Evidence from Antidumping Policy — **WP-09-17**
Chad P. Bown and Meredith A. Crowley

Too much right can make a wrong: Setting the stage for the financial crisis — **WP-09-18**
Richard J. Rosen

Can Structural Small Open Economy Models Account for the Influence of Foreign Disturbances? — **WP-09-19**
Alejandro Justiniano and Bruce Preston

Liquidity Constraints of the Middle Class — **WP-09-20**
Jeffrey R. Campbell and Zvi Hercowitz

Monetary Policy and Uncertainty in an Empirical Small Open Economy Model — **WP-09-21**
Alejandro Justiniano and Bruce Preston

Firm boundaries and buyer-supplier match in market transaction: IT system procurement of U.S. credit unions — **WP-09-22**
Yukako Ono and Junichi Suzuki

Health and the Savings of Insured Versus Uninsured, Working-Age Households in the U.S. — **WP-09-23**
Maude Toussaint-Comeau and Jonathan Hartley

The Economics of "Radiator Springs:" Industry Dynamics, Sunk Costs, and Spatial Demand Shifts — **WP-09-24**
Jeffrey R. Campbell and Thomas N. Hubbard

On the Relationship between Mobility, Population Growth, and Capital Spending in the United States — **WP-09-25**
Marco Bassetto and Leslie McGranahan

The Impact of Rosenwald Schools on Black Achievement — **WP-09-26**
Daniel Aaronson and Bhashkar Mazumder

Working Paper Series *(continued)*

Working Paper Series *(continued)*

The Role of Housing in Labor Reallocation **WP-10-18**
Morris Davis, Jonas Fisher, and Marcelo Veracierto

Why Do Banks Reward their Customers to Use their Credit Cards? **WP-10-19**
Sumit Agarwal, Sujit Chakravorti, and Anna Lunn

The impact of the originate-to-distribute model on banks
before and during the financial crisis **WP-10-20**
Richard J. Rosen

Simple Markov-Perfect Industry Dynamics **WP-10-21**
Jaap H. Abbring, Jeffrey R. Campbell, and Nan Yang

Commodity Money with Frequent Search **WP-10-22**
Ezra Oberfield and Nicholas Trachter

Corporate Average Fuel Economy Standards and the Market for New Vehicles **WP-11-01**
Thomas Klier and Joshua Linn

The Role of Securitization in Mortgage Renegotiation **WP-11-02**
Sumit Agarwal, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet,
and Douglas D. Evanoff

Market-Based Loss Mitigation Practices for Troubled Mortgages
Following the Financial Crisis **WP-11-03**
Sumit Agarwal, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet,
and Douglas D. Evanoff

Federal Reserve Policies and Financial Market Conditions During the Crisis **WP-11-04**
Scott A. Brave and Hesna Genay

The Financial Labor Supply Accelerator **WP-11-05**
Jeffrey R. Campbell and Zvi Hercowitz

Survival and long-run dynamics with heterogeneous beliefs under recursive preferences **WP-11-06**
Jaroslav Borovička

A Leverage-based Model of Speculative Bubbles (Revised) **WP-11-07**
Gadi Barlevy

Estimation of Panel Data Regression Models with Two-Sided Censoring or Truncation **WP-11-08**
Sule Alan, Bo E. Honoré, Luojia Hu, and Søren Leth–Petersen

Fertility Transitions Along the Extensive and Intensive Margins **WP-11-09**
Daniel Aaronson, Fabian Lange, and Bhashkar Mazumder

Black-White Differences in Intergenerational Economic Mobility in the US **WP-11-10**
Bhashkar Mazumder

Can Standard Preferences Explain the Prices of Out-of-the-Money S&P 500 Put Options? **WP-11-11**
Luca Benzoni, Pierre Collin-Dufresne, and Robert S. Goldstein

Working Paper Series *(continued)*

Business Networks, Production Chains, and Productivity:
A Theory of Input-Output Architecture **WP-11-12**
Ezra Oberfield