A Theory of Input-Output Architecture*

Ezra Oberfield

Princeton University

ezraoberfield@gmail.com

February 16, 2017

Abstract

Individual producers exhibit enormous heterogeneity in many dimensions. This paper develops a theory in which the network structure of production—who buys inputs from whom—forms endogenously. Entrepreneurs produce using labor and exactly one intermediate input; the key decision is which other entrepreneur's good to use as an input. Their choices collectively determine the economy's equilibrium input-output structure, generating large differences in size and shaping both individual and aggregate productivity. When the elasticity of output to intermediate inputs in production is high, star suppliers emerge endogenously. This raises aggregate productivity as, in equilibrium, more supply chains are routed through higher-productivity techniques.

Keywords: Networks, Productivity, Supply Chains, Size distribution, Input-Output Structure, Economic Growth, Ideas

JEL Codes: O31, O33, O47, L11, L25

*This paper previously circulated as "Business Networks, Production Chains, and Productivity: A Theory of Input-Output Architecture." I am grateful for comments from Amanda Agan, Fernando Alvarez, Enghin Atalay, Gadi Barley, Macro Basset, Jared Boracic, Parco Bureau, Jeff Campbell, Vasto Carvalho, Thomas Chaney, Meredith Crowley, Cristina De Nardi, Aspen Gorry, Joe Kaboski, Sam Kortum, Alejandro Justiniano, Robert Lucas, Devesh Raval, Richard Rogerson, Rob Shimer, Nancy Stokey, Nico Trachter, Mark Wright, Andy Zuppann, and various seminar participants. I also thank Todd Messer, Matt Olson, and Kazu Matsuda for excellent research assistance. All mistakes are my own.

> "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."
>
> Rosenberg (1979), p. 28-29

One of the more striking features of microeconomic production data is the enormous heterogeneity across producers. Researchers have documented vast differences in employment, sales, and, more recently, engagement in input-output linkages.[1] These observations have generated interest in two key questions: Why is there so much microeconomic heterogeneity? And what are the implications of this heterogeneity for aggregate productivity?

This paper develops a theory in which the economy's input-output architecture arises endogenously and shows that enormous differences in size can emerge even when differences in productivity are arbitrarily small. The theory offers a new perspective relative to canonical models of firm dynamics which typically explain dispersion in size as resulting from the accumulation of random growth of productivity or demand residuals.[2]

The theory is based on the premise that there may be multiple ways to produce a good, each with a different set of inputs. In the model, each entrepreneur sells a particular good and can use a variety of techniques to produce that good. Each technique allows the entrepreneur to produce her good using labor and exactly one other entrepreneur's good as an intermediate input, with productivity specific to that input. The key decision is which other entrepreneur's good to use. An entrepreneur's cost of production when using one of these techniques depends both on the technique's productivity and on price of the intermediate input which, in turn, depends on the production cost of the entrepreneur that produces that input. In this environment, the economy's production possibilities cannot be summarized in terms of the capabilities of individual producers. The collection of known production techniques forms a network comprising each entrepreneur's potential suppliers and potential customers—others who might use the entrepreneur's good as an intermediate input. When producing, each entrepreneur selects from her techniques the one that is most cost-effective. These choices

[1] See, for example, Axtell (2001), Rossi-Hansberg and Wright (2007), Acemoglu et al. (2012), and Atalay et al. (2011).

[2] These include Gibrat (1931), Simon and Bonini (1958), Hopenhayn (1992), Klette and Kortum (2004), Rossi-Hansberg and Wright (2007), and Luttmer (2007).

collectively determine each producer's size and contribution to aggregate productivity.

Terms of trade determine which of the potential input-output linkages are used in equilibrium and the extent to which one producer's low cost of production is passed on to others. Producers engage in standard monopolistic competition in sales of their goods to a representative household for consumption, but bilateral contracts govern each pairwise transaction. I restrict attention to arrangements that are countably stable; terms of trade must be such that countable coalitions cannot find alternative terms that are mutually beneficial.[3] In any such equilibrium, production within each supply chain is efficient.

The paper's main results describe how individual choices lead to the endogenous emergence of star suppliers and the implications for aggregate outcomes.[4] Star suppliers are individual entrepreneurs that, in equilibrium, sell their goods to many other entrepreneurs for intermediate use.[5] For an entrepreneur to be a star supplier, she must have many potential customers, and a large fraction of those potential customers must choose to use her good in equilibrium. Whether those potential customers choose her good depends both on the price she is willing to accept for her good and on how much those potential customers are willing to trade off a technique with a supplier that offers a low price for a technique with a high match-specific productivity.[6]

One striking feature of this environment is that the prevalence of star suppliers—and hence dispersion in size—is independent of the parameter that determines dispersion in entrepreneurs' marginal cost. Thus even with arbitrarily little variation in entrepreneurs' marginal costs, there can still be large differences in size.

To shed light on the economic forces shaping the organization of production and ag-

[3]I also discuss the implications of a weaker restriction, pairwise stability. Pairwise stable arrangements are not robust to deviations that would be natural in this environment. For example, if an entrepreneur agrees to terms with a new supplier that would lower her marginal cost, it would be natural for that entrepreneur to reoptimize terms of trade with any buyers (and for those buyers to reoptimize terms of trade with their buyers, etc.), and for the new supplier to alter the quantity of inputs she purchases from her supplier, etc.

[4]The model is constructed to speak to patterns of linkages among individual producers rather than broader patterns of which sectors buy inputs from which other sectors. See Carvalho and Voigtländer (2014) for a promising step in that direction.

[5]Building on Gabaix (2011), Acemoglu et al. (2012) show that the prevalence of star suppliers is one factor that determines whether idiosyncratic shocks are relevant for aggregate fluctuations.

[6]Since Rosen (1981), the superstar literature studies how small differences in talent can lead to large differences in compensation. The key factor is the relationship between differences in talent and how much customers are willing to trade off differences in talent for a lower price. While the analogy is not perfect, similar forces are at play here.

gregate productivity, I impose a functional form assumption that allows for an analytical characterization of a number of features of the economy. The key part of the assumption is that the distribution from which techniques' productivities are drawn has a right tail that follows a power law with exponent ζ. Under this assumption, I derive four results about the way entrepreneurs selections of suppliers interact to determine the size distribution, matching patterns, expenditure shares, and aggregate productivity.

I first study which features of the environment determine how customers are distributed across entrepreneurs and the size distribution. The distribution of customers depends on a single parameter, α; each technique is a Cobb-Douglas production function, and α is the elasticity of output of the buyer's good with respect to the supplier's good. The prevalence of star suppliers depends on the right tail of this distribution. I show that the distribution has a power-law tail with exponent $1/\alpha$. Why does α play the key role? Recall that when an entrepreneur selects a supplier, she considers both the match-specific productivity and the cost of the associated input of each of her techniques. When α is small, the cost of the inputs is less important, and thus suppliers' cost are less important drivers of choices of suppliers. Conversely, when α is large, entrepreneurs with low production costs are selected as suppliers more systematically, and are therefore more likely to be star suppliers. In line with the results mentioned above, the distribution does not depend on ζ, which indexes dispersion in marginal costs.

The distribution of customers across entrepreneurs is one determinant of the size distribution. An entrepreneur's size, as measured by employment, depends on her sales to the household and her sales to each customer, which depends on the size of those customers. It turns out that, under the functional form assumption, the size distribution has a threshold property. If α is small, the model has the property found in many other models that the right tail is dominated by those that sell large quantities to the household, and the shape of the right tail depends on dispersion of marginal cost (ζ). If α is larger, intermediate inputs become more important in production and sales of intermediate inputs become more concentrated in star suppliers. If α is large enough to cross a threshold, the right tail becomes dominated by star suppliers and follows a power law with exponent $1/\alpha$. Thus the extent of heterogeneity in productivity or in marginal cost no longer plays any role in determining the shape of the right tail.

Second, I show that one's conclusions about whether the equilibrium features assortative matching depends on which attribute one focuses on. Among equilibrium matches, buyers' and suppliers' marginal costs are uncorrelated but their sizes (as measured by employment) are positively correlated. For a randomly selected technique, if the potential supplier has a low cost of production the technique is likely to deliver a low cost of production to the potential buyer. In equilibrium, however, techniques are not selected randomly. An entrepreneur is likely to select a technique whose supplier has a low marginal cost even if the match-specific productivity is poor, but is unlikely to select a technique whose supplier has a high marginal cost unless the match-specific productivity is especially high. As a result, among matches observed in equilibrium, buyers' and suppliers' marginal costs are uncorrelated.[7] Despite this, their sizes are positively correlated; a supplier whose customer is unusually large needs to hire an unusually large amount of labor to produce the intermediates for that customer.

Third, while dispersion of marginal cost has no impact on the prevalence of star suppliers, it does determine the aggregate cost share of intermediate inputs. In a competitive equilibrium, the cost share of intermediate inputs would be α. Here, however, since buyers and suppliers split surplus, the cost share is weakly larger.[8] While countable-stability does not pin down how producers split surplus, there is a class of equilibria indexed by a parameter that has a natural interpretation as bargaining power.[9] Notably, within this class the aggregate cost share depends on ζ, which indexes dispersion in marginal cost and is related to the size of the loss when an entrepreneur must switch to its next best supply chain.

Finally, the emergence of star suppliers affects aggregate productivity in the following

[7]One of the most promising applications of data on firm-to-firm transactions is using the observed network structure to learn about the production process—who supplies inputs to whom and how these interactions change over time. Since much of the literature is focused on the source of fluctuations, an effort has been made to identify shocks and decompose the fluctuations (Foerster et al. (2011), Atalay (2013), Di Giovanni et al. (2014), Kramarz et al. (2016)). An important lesson from the labor literature (Abowd et al. (1999), Lopes de Melo (Forthcoming)) is that ignoring the endogeneity of match formation can bias inference about individual characteristics. The fact that individuals select into a match provides information about those individuals and about the quality of the match. The same biases arise in the environment presented here.

[8]Thus in contrast to environments such as Hulten (1978) in which producers are price takers, the cost share of intermediate inputs is not a sufficient statistic for the input-output multiplier.

[9]The bargaining power parameter is analogous to that of the generalized Nash bargaining solution. The difference here is that the size of surplus to be split between a supplier and buyer depends on how the buyer splits surplus with *her* buyers, etc.

sense: entrepreneurs' selections jointly determine the supply chains used to produce each good. Aggregate productivity depends on the match-specific productivity of the techniques used at each step in each of those supply chains. When α is larger, supply chains are more likely to be routed through the most productive techniques in the economy, raising aggregate productivity. This channel complements the usual input-output multiplier that is present in all models with roundabout production.[10]

Related Literature

The model provides a mechanism that delivers a skewed cross-sectional distribution of links which complements but differs from that of the influential preferential attachment model of Barabasi and Albert (1999).[11] Here, even though the distribution of potential customers follows a distribution with a thin tail, the distribution of actual customers endogenously exhibits more variation. The mechanism is a natural one once choices are endogenous: entrepreneurs that are willing to accept a lower price for their goods are systematically more likely to be selected as suppliers by their potential customers.

Similarly, if α is large enough, the model's predictions for the size distribution differ substantively from canonical models of producer dynamics referenced in footnote 2. In those models, dispersion in producer size results from the accumulation of idiosyncratic productivity or demand shocks. The model here is simple in the sense that there is a single type of shock: the arrival of a technique. But there are two key differences from productivity

[10]By now it is well understood that input-output linkages play a key role in shaping how one producer's cost of production impacts other producers and ultimately aggregate productivity. The literature on aggregate fluctuations has focused on how sectoral shocks propagate through an economy – see Long and Plosser (1983), Horvath (1998), Dupor (1999), Acemoglu et al. (2012), and Carvalho and Gabaix (2013). In the growth literature, see Jones (2011, 2013), and Ciccone (2002). In the trade literature, Caliendo and Parro (2011) argue that the welfare impact of changes in tariffs in particular industries depends on the input-output structure.

[11]The preferential attachment model was designed because network models with uniform arrival of links exhibit link distributions with thin tails, whereas many real world networks feature link distributions with power-law tails. In the preferential attachment model, there is an initial network and, over time, new links are formed and new nodes are born. The probability that a new link involves a particular node is increasing in the number of links that node already has. Variants of the preferential attachment model of Barabasi and Albert (1999) have been used to explain the cross-sectional distribution of the firm to firm linkages (Atalay et al. (2011)) and the network structure of international shipments of goods (Chaney (2011)). See also Kelly et al. (2013). The mechanism is similar to proportional random growth models and goes back to Yule (1925) and Simon (1955).

shocks of canonical models. First, a technique represents both a way to produce one good but also a way of using a different good.[12] In other words, the arrival of a technique is a supply shock to one producer (the buyer) but a demand shock to another producer (the supplier). Second, a producer is affected by shocks to others (i.e., to customers' customers or to suppliers' suppliers, etc.). It is the interaction of upstream and downstream techniques and the higher-order interconnections that generates the rich cross-sectional patterns.

The model is written in such a way that it nests a simple version of Kortum (1997) when α is zero—a special case in which the network structure plays no role.[13] This special case provides a backdrop against which one can see the network structure's role in shaping economic outcomes.

Finally, the market structure builds on Hatfield et al. (2013) who study stable outcomes in supply chain networks. The main substantive difference is that I focus on countably-stable outcomes and splits of surplus consistent with bargaining power whereas Hatfield et al. (2013) focus on fully stable outcomes and compare them to competitive equilibria. The main technical difference is that the environment studied here contains a continuum of agents, which allows the use of the law of large numbers but raises some technical issues such as defining feasible outcomes.[14]

1 The Environment

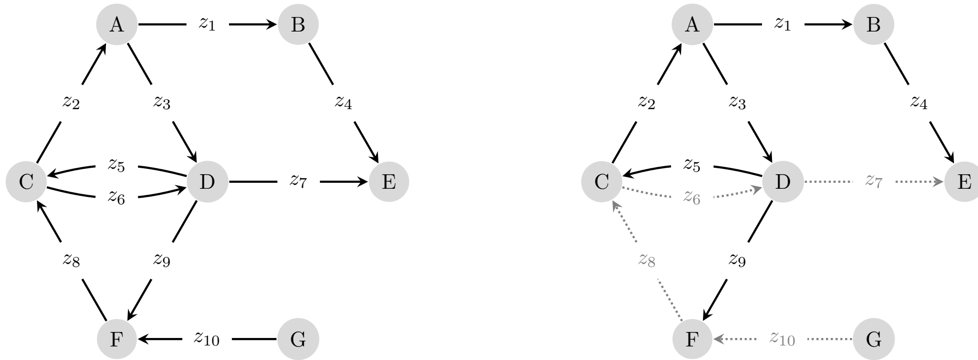
There is a unit mass of entrepreneurs, and each is the sole producer of a differentiated good. A representative household derives utility from consuming all of the goods, consuming the bundle $Y^0 = \left(\int_0^1 (y_j^0)^{\frac{\varepsilon-1}{\varepsilon}} dj\right)^{\frac{\varepsilon}{\varepsilon-1}}$ where y_j^0 is consumption of good j, and supplies L units of labor inelastically.

To produce, an entrepreneur must use a technique. A technique is a production func-

[12]The idea that innovation is finding a new use for an existing good is related to recombinant growth of Weitzman (1998), in which innovation is finding a better way to combine inputs. Scherer (1982) and Pavitt (1984) show that roughly three quarters of innovations are for use by others outside the innovating firm's sector.

[13]The model is also related to Alvarez et al. (2008), Lucas (2009), Lucas and Moll (2014) and Perla and Tonetti (2014). See Buera and Oberfield (2016) for an elaboration of how the mathematical structure of this model is related to models of idea flows.

[14]Countable stability is also related to f-core of Hammond et al. (1989) who studied deviations of finite coalitions in a continuum economy.



(a) The Set of Techniques

(b) Techniques that are Used

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

Figure 1a gives an example of a set of techniques, Φ. Each node is an entrepreneur and edges correspond to techniques. An edge's direction indicates which entrepreneur produces the output and which provides the input. An edge's number represents the productivity of the technique. In Figure 1b, solid arrows are techniques that are used.

tion with which one entrepreneur can produce her good using labor and exactly one other entrepreneur's good as an intermediate input. We denote a single technique by ϕ and the set of all techniques by Φ. A technique ϕ is specific to a buyer, denoted $b(\phi)$, specific to a supplier, $s(\phi)$, and has a match-specific productivity $z(\phi)$. A technique is fully defined by these three characteristics. The technique describes the production function

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

where y units of the buyer's good is produced using x units of the supplier's good and l units of labor.

The economy's production possibilities can be completely described by the set of all the techniques available to the different entrepreneurs, Φ. Techniques arrive randomly via a process described below. Any realization of the set of techniques, Φ, can be represented by a weighted, directed graph. Figure 1a gives an example of a realization for an economy with a finite number of entrepreneurs. In the figure, each entrepreneur is a node and each technique is a edge connecting two nodes. Each edge has a direction, indicating which entrepreneur would supply the intermediate input and which would produce the output, and a number corresponding to the technique's match-specific productivity.

Some entrepreneurs may have multiple ways to produce whereas others may have none. For the realization of Φ depicted in Figure 1a, it is infeasible for G to produce, but E has a technique to produce using good D as an input and one that uses good B. In principle she may produce using both, although it will be shown later that in equilibrium she will generically produce using only one. While Figure 1a completely describes the economy's production possibilities, Figure 1b gives an example of entrepreneurs' selections of which techniques to use. Such selections, which ultimately depend on prices, correspond to vertical relationships we would observe in equilibrium and jointly determine the supply chains that are used to produce each good.

The set of techniques Φ is a random set. The number of techniques with which an entrepreneur can produce her good follows a Poisson distribution with mean M. For each of those techniques, the identity of the supplier $s(\phi)$ is random and uniformly drawn from all entrepreneurs in the economy. This incidentally implies that the number of techniques for which an entrepreneur is the supplier also follows a Poison distribution with mean M. Each technique's match-specific productivity, $z(\phi)$, is drawn from a fixed distribution with CDF H.[15] Assumption 1 guarantees that the right tail of this distribution is not too thick, which is sufficient to ensure that aggregate output is almost surely finite.

Assumption 1. *The support of H is bounded below by some $z_0 > 0$ and there exists a $\beta > \varepsilon - 1$ such that $\lim_{z\to\infty} z^{\beta}[1 - H(z)] = 0$.*

Φ implicitly determines the supply chains that may be used to produce each good. For an entrepreneur, a supply chain consists of a technique she can use produce her good, a technique that the associated supplier can use, etc. Formally, a *supply chain* for entrepreneur j is an infinite sequence of techniques $\{\phi_k\}_{k=0}^{\infty}$ with the property that $j = b(\phi_0)$ and $s(\phi_k) = b(\phi_{k+1})$ for each k. Supply chains may cycle, in which case the techniques in the sequence repeat.[16] It is feasible for an entrepreneur to produce only if there is at least one supply chain to

[15] To keep the notation manageable, I abstract from any ex-ante differences in goods' suitability for use as an intermediate input for any other type of good or for consumption. In many applications, ex-ante asymmetries across goods would be important (reflecting different industries or countries). The model can easily accommodate such asymmetries without loss of tractability; see working paper version.

[16] For example, lumber is useful in the production of an ax, while an ax is useful in the production of lumber.

produce that entrepreneur's good.[17][18]

2 Market Structure

Terms of trade among entrepreneurs determine entrepreneurs' choices of inputs, productions decisions, and productivity. This section describes a market structure and the terms of trade that arise in equilibrium. The market structure is motivated by the idea that groups of entrepreneurs that are small (relative to the market) may choose terms that maximize their mutual gains from trade.

Entrepreneurs engage in monopolistic competition when selling to the representative household, but sales of goods for intermediate use are governed by bilateral trading contracts specifying a buyer, a supplier, a quantity of the supplier's good to be sold to the buyer and a payment. Given a contracting arrangement, each entrepreneur makes her remaining production decisions to maximize profit. Goods may not be resold and entrepreneurs remit all profit to the household.

The economy is in equilibrium when the arrangement is such that no coalition of entrepreneurs (of a specified size) would find it mutually beneficial to deviate by altering terms of trade among members of the coalition and/or dropping contracts with those not in the coalition.

2.1 Arrangements and Payoffs

Terms of trade are described by a contracting arrangement. An *arrangement* is a double $\{x(\phi), T(\phi)\}$ for each technique $\phi \in \Phi$ where $x(\phi)$ indicates the quantity of good $s(\phi)$ to be sold to the buyer $b(\phi)$, and $T(\phi)$ indicates a payment to be made from the buyer to the supplier.[19]

[17]In the example with a finite set of entrepreneurs, any supply chain must contain a cycle, With a continuum of entrepreneurs, a supply chain need not cycle. Further, the set of entrepreneurs who have supply chains that cycle has measure zero.

[18]The assumptions that production of each good requires the use of some other good as an input and that chains of these relationships may continue indefinitely follow in the tradition of Leontief (1951) and other models with roundabout production.

[19]The notation rules out the possibility of a contract between a buyer and supplier when the buyer does not have a technique that uses the supplier's good as an input. Online Appendix A.1 shows that this is

Taking as given the arrangement, the wage, and the household's demand for its good, each entrepreneur makes production decisions to maximize profit. These production decisions consist of how much output to produce using each of her available techniques, how much labor to hire, and a price at which output is sold to the household. If entrepreneur j uses $l(\phi)$ units of labor with technique ϕ, then its output of good j produced using that technique is $y(\phi) \equiv \frac{z(\phi)x(\phi)^{\alpha}l(\phi)^{1-\alpha}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}$.

For entrepreneur j, there are two sets of techniques that are directly relevant. Let $U_j \equiv \{\phi \in \Phi : b(\phi) = j\}$ be the techniques for which j is the buyer (those upstream from j) and let $D_j \equiv \{\phi \in \Phi : s(\phi) = j\}$ be the techniques for which j is the supplier (those downstream from j). These respectively determine j's potential suppliers and potential customers.

Total output of good j and total labor usage by j are thus $y_j \equiv \sum_{\phi \in U_j} y(\phi)$ and $l_j \equiv \sum_{\phi \in U_j} l(\phi)$ respectively. The total usage of good j is output for consumption and for use as an intermediate input by others, $y_j^0 + \sum_{\phi \in D_j} x(\phi)$.

Given the arrangement, entrepreneur j's payoff is

$$\pi_j = \max_{p_j^0, y_j^0, \{l(\phi)\}_{\phi \in U_j}} p_j^0 y_j^0 + \sum_{\phi \in D_j} T(\phi) - \sum_{\phi \in U_j} [T(\phi) + wl(\phi)] \tag{1}$$

subject to satisfying the household's demand

$$y_j^0 \leq Y^0 (p_j^0/P^0)^{-\varepsilon}, \qquad P^0 \equiv \left(\int_0^1 (p_j^0)^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}}$$

and a technological constraint that total usage of good j cannot exceed total production of good j

$$y_j^0 + \sum_{\phi \in D_j} x(\phi) \leq \sum_{\phi \in U_j} \frac{z(\phi)x(\phi)^{\alpha}l(\phi)^{1-\alpha}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}} \tag{2}$$

without loss of generality. Because goods may not be resold, such a transaction can never be optimal; a buyer would drop any contract with a positive payment whereas a supplier would drop all other contracts. Online Appendix A.1 shows that such transactions are also not essential off the equilibrium path. The notation also rules out multiple contracts for the same technique. Online Appendix A.1 shows that this restriction can be made without loss of generality as well.

2.2 Feasibility

An allocation consists of quantities of labor and the intermediate input used for and output from each technique $\{l(\phi), x(\phi), y(\phi)\}_{\phi\in\Phi}$, and consumption of the household $\{y_j^0\}_{j\in[0,1]}$. An allocation is *feasible* if it is both *resource feasible* and *chain feasible*. An allocation is *resource feasible* if it satisfies (2) for each j and the labor resource constraint,

$$\int_0^1 \left(\sum_{\phi \in U_j} l(\phi) \right) dj \leq L$$

We impose a second requirement, chain feasibility, which is analogous to a no-Ponzi condition. For entrepreneur j, let $\mathfrak{C}_j$ denote the set of supply chains available to produce good j. Loosely, for a supply chain $\mathfrak{c} \in \mathfrak{C}_j$ available to entrepreneur j, chain stability requires that the final output of good j produced using chain $\mathfrak{c}$ is feasible given the inputs used at each step in the supply chain. Defining chain feasibility requires an alternative representation of the allocation that we label a supply chain representation. This is a decomposition of the allocation into production within the many supply chains available to produce the various goods. For each chain $\mathfrak{c}$ and $n \in \{0, 1, 2, ...\}$, the variables $\{y^n(\mathfrak{c}), x^n(\mathfrak{c}), l^n(\mathfrak{c})\}$ are the quantities of output, intermediate input, and labor used in the n^{th}-to-last step in the supply chain $\mathfrak{c}$ in the production of good j for consumption.[20] A supply chain representation must satisfy three properties. First, for any chain $\mathfrak{c}$ and n:

$$y^n(\mathfrak{c}) = \frac{1}{\alpha^\alpha (1-\alpha)^{1-\alpha}} z^n(\mathfrak{c}) x^n(\mathfrak{c})^\alpha l^n(\mathfrak{c})^{1-\alpha} \tag{3}$$

where $z^n(\mathfrak{c})$ is the match-specific productivity of the n^{th}-to-last technique in the supply chain. Second, the output at one stage is the intermediate input used at the next stage, $x^n(\mathfrak{c}) = y^{n+1}(\mathfrak{c})$. Third, the supply chain representation must actually match the allocation.[21]

[20]Supply chains overlap. To avoid double counting, the decomposition specifies production of supply chain $\mathfrak{c} \in \mathfrak{C}_j$ as the production of good j for consumption rather than total output of good j, as the latter includes production of good j for intermediate use in other supply chains. Since every technique exhibits constant returns to scale, the decomposition is well-defined.

[21]Formally, if $\mathfrak{C}(n, \phi)$ is the set of chains for which the technique ϕ is the n-th most downstream technique, then $l(\phi) = \sum_{n=0}^{\infty} \sum_{\mathfrak{c}\in\mathfrak{C}(n,\phi)} l^n(\mathfrak{c})$, $x(\phi) = \sum_{n=0}^{\infty} \sum_{\mathfrak{c}\in\mathfrak{C}(n,\phi)} x^n(\mathfrak{c})$, and $y(\phi) = \sum_{n=0}^{\infty} \sum_{\mathfrak{c}\in\mathfrak{C}(n,\phi)} y^n(\mathfrak{c})$.

Online Appendix A.2 provides a more detailed treatment.[22]

An allocation is *chain feasible* if there exists a supply chain representation such that for every chain,

$$y^0(\mathfrak{c}) \leq \prod_{n=0}^{\infty} \left(\frac{1}{\alpha^\alpha (1-\alpha)^{1-\alpha}} z^n(\mathfrak{c}) l^n(\mathfrak{c})^{1-\alpha} \right)^{\alpha^n}$$

To unpack this definition, note that combining (3) for stages $n = 0, ..., k$ for a single chain $\mathfrak{c}$ and using $x^n(\mathfrak{c}) = y^{n+1}(\mathfrak{c})$ yields $y^0(\mathfrak{c}) = x^k(\mathfrak{c})^{\alpha^{k+1}} \prod_{n=0}^{k} \left(\frac{1}{\alpha^\alpha(1-\alpha)^{1-\alpha}} z^n(\mathfrak{c}) l^n(\mathfrak{c})^{1-\alpha} \right)^{\alpha^n}$. Chain feasibility imposes that the sequence of intermediate inputs used further and further upstream in the supply chain does not explode, i.e., $\limsup_{k\to\infty} \left[x^k(\mathfrak{c})\right]^{\alpha^{k+1}} = 1$.

2.3 Stability and Equilibrium

A coalition is a set of entrepreneurs, J. For a coalition, a *deviation* consists of (i) dropping any subset of contracts that involve at least one entrepreneur in J–setting both the quantity and the payment specified in the contract to zero; and (ii) altering the terms of any contract for which both the buyer and supplier are members of the coalition.

A *dominating deviation* for a coalition is a deviation that would deliver weakly higher payoffs to all members of the coalition and strictly higher payoffs to at least one member of the coalition.

An arrangement is *n-stable* if there are no dominating deviations for coalitions of size/cardinalty n or smaller. The text below will discuss both pairwise-stable arrangements and countably-stable arrangements—arrangements for which no countable coalition has a dominating deviation. Of course, every countably-stable arrangement is pairwise stable, but not vice versa.

So that payoffs are well-defined out of equilibrium, we must account for an arrangement that dictates that an entrepreneur is obliged to produce ($\sum_{\phi\in D_j} x(\phi) > 0$) but receives no intermediate inputs with which to produce ($\sum_{\phi\in U_j} x(\phi) = 0$). To this end, I adopt the convention that, in such a situation, the entrepreneur can meet its obligation at infinite cost.

Finally, we define an n-stable equilibrium. An *n-stable equilibrium* is an arrangement

[22] If the allocation is such that each entrepreneur produces using a single technique, there is a unique supply chain representation; in instances in which at least one entrepreneur produces using multiple techniques, there are multiple supply chain representations. Nevertheless, for any supply chain representation, there is a unique allocation.

$\{x(\phi), T(\phi)\}_{\phi\in\Phi}$, entrepreneurs' choices, $\left\{p_j^0, y_j^0, \{l(\phi)\}_{\phi\in U_j}\right\}_{j\in[0,1]}$, and a wage w such that (i) Given the wage, total profit, and prices, the consumption choices $\{y_j^0\}_{j\in[0,1]}$ maximize the representative household's utility; (ii) For each $j \in [0,1]$, $\left\{p_j^0, y_j^0, \{l(\phi)\}_{\phi\in U_j}\right\}$ maximize j's payoff given the arrangement, the wage, and the household's demand; (iii) Labor and final goods markets clear; (iv) There are no coalitions of size/cardinality n with dominating deviations; (v) The allocation is feasible.

2.4 Properties of Equilibria

Let λ_j be entrepreneur j's marginal cost, the multiplier on (2).[23] For any arrangement, j will choose to sell her good to the household at the usual markup over marginal cost, $p_j^0 = \frac{\varepsilon}{\varepsilon-1}\lambda_j$ while her expenditure on labor to use with technique $\phi \in U_j$ is $wl(\phi) = (1-\alpha)\lambda_j y(\phi)$.

It will be convenient to define $q_j \equiv \frac{w}{\lambda_j}$ to be the inverse of j's marginal cost in units of labor. q_j is the efficiency with which good j can be produced, or for shorthand, j's *efficiency*. Let $Q \equiv \left(\int_0^1 q_j^{\varepsilon-1} dj\right)^{\frac{1}{\varepsilon-1}}$ be the usual Dixit-Stiglitz productivity aggregator of entrepreneurs' efficiencies. Proposition 1 characterizes implications of both pairwise stability and countable stability and shows that Q is a measure of aggregate productivity for the economy.

Proposition 1. *In any pairwise-stable equilibrium, entrepreneur j sells $y_j^0 = q_j^\varepsilon Q^{1-\varepsilon} L$ units of her good to the household for consumption, the household's consumption aggregate is*

$$Y^0 = QL \tag{4}$$

and entrepreneurs' efficiencies satisfy

$$q_j = \max_{\phi\in U_j} z(\phi) q_{s(\phi)}^\alpha, \qquad \text{or } q_j = 0 \text{ if } U_j \text{ is empty} \tag{5}$$

$$q_j \leq \sup_{\mathfrak{c}\in\mathfrak{C}_j} \prod_{n=0}^{\infty} \left[z^n(\mathfrak{c})\right]^{\alpha^n}, \qquad \text{or } q_j = 0 \text{ if } \mathfrak{C}_j \text{ is empty} \tag{6}$$

[23] Note that $\lambda_j = \infty$ if either U_j is empty or j has no intermediate inputs (i.e., $\sum_{\phi\in U_j} x(\phi) = 0$).

In any countably-stable equilibrium,

$$q_j = \sup_{\mathfrak{c} \in \mathfrak{C}_j} \prod_{n=0}^{\infty} [z^n(\mathfrak{c})]^{\alpha^n}, \qquad \text{or } q_j = 0 \text{ if } \mathfrak{C}_j \text{ is empty} \tag{7}$$

Every countably-stable equilibrium is efficient. If H is atomless then with probability one there is an essentially unique allocation consistent with countable stability.

(5) summarizes how pairwise stability connects each entrepreneur's marginal cost to that of its suppliers. The proof shows that, for every technique,

$$\lambda_{b(\phi)} \leq \frac{1}{z(\phi)} \lambda^{\alpha}_{s(\phi)} w^{1-\alpha}, \qquad \text{with equality if } x(\phi) > 0 \tag{8}$$

To understand this, note that for any technique, the buyer's marginal cost of output from using that technique is $\frac{1}{z(\phi)} \chi^{\alpha}_{s(\phi)} w^{1-\alpha}$ where $\chi_{s(\phi)}$ is the buyer's shadow value of the supplier's good. The key implication of pairwise stability is that for each technique that is actually used, (i.e., $x(\phi) > 0$), the buyer's shadow value of an input must equal the supplier's marginal cost, as this maximizes the bilateral gains from trade.[24] For entrepreneur j, collecting these conditions for all techniques in U_j and using $q_j = w/\lambda_j$ gives the result.

Along similar lines, (4) says that in any pairwise stable equilibrium aggregate productivity is simply the usual Dixit-Stiglitz productivity aggregator of entrepreneurs' efficiencies. The key step in the proof is to show pairwise stability implies that each entrepreneur's marginal cost—measured in units of labor—is equal to the quantity of labor used across all stages in all supply chains to produce a unit of that entrepreneur's good. In other words, pairwise stability is sufficient to ensure that no supply chain suffers from double marginalization; each pair maximizes their bilateral gains from trade, which ensures that the quantity in the contract is such that the supplier's marginal cost is equal to the buyer's shadow value of the input.

(6) is an implication of chain feasibility. For a single chain, $\mathfrak{c} \in \mathfrak{C}_j$, it follows from combining (5) for entrepreneur j, its suppliers, its suppliers' suppliers, etc. and imposing

[24] If $\lambda_{b(\phi)} < \frac{1}{z(\phi)} \lambda^{\alpha}_{s(\phi)} w^{1-\alpha}$, the buyer and supplier could increase their bilateral surplus by lowering $x(\phi)$. Such a deviation is always possible unless $x(\phi) = 0$.

chain feasibility.[25]

(7) shows that if an arrangement is countably stable, entrepreneur j's efficiency equals that of its most efficient supply chain.[26] The proof shows that if there is an entrepreneur j with $q_j < \sup_{\mathfrak{c}\in\mathfrak{C}_j} \prod_{n=0}^{\infty} [z^n(\mathfrak{c})]^{\alpha^n}$ then there is a countable coalition made up of the entrepreneurs that form some chain $\mathfrak{c} \in \mathfrak{C}_j$ with $\prod_{n=0}^{\infty} [z^n(\mathfrak{c})]^{\alpha^n} > q_j$ with a dominating deviation in which each of those entrepreneurs produces slightly more. This would raise j's profit enough to compensate all of those entrepreneurs in the chain.

The proposition states that every countably-stable equilibrium is efficient. The proof shows that if a planner controlled production throughout the entire supply chain $\mathfrak{c} \in \mathfrak{C}_j$ and allocated some quantity of labor l optimally throughout the chain, then it would solve $y^0(\mathfrak{c}) = \max_{\{l^n(\mathfrak{c})\}} \prod_{n=0}^{\infty} \left(\frac{1}{\alpha^\alpha(1-\alpha)^{1-\alpha}} z^n(\mathfrak{c}) l^n(\mathfrak{c})^{1-\alpha}\right)^{\alpha^n}$ subject to $\sum l^n(\mathfrak{c}) \leq l$. With the optimal allocation of labor across stages, the planner would face the indirect production function $y^0(\mathfrak{c}) = \prod_{n=0}^{\infty} [z^n(\mathfrak{c})]^{\alpha^n} l$. Thus in every equilibrium, each entrepreneur's marginal cost matches the planner's marginal shadow cost of the good.[27]

Proposition 1 also states that, generically, the allocation of goods and labor is essentially unique. The exception is if some entrepreneur j has access to two techniques would deliver efficiency q_j. In that case countable stability does not determine how the entrepreneur splits her production across the two techniques. Given the stochastic process of the arrival of techniques, however, the probability that any entrepreneur has two techniques that deliver exactly the same efficiency is zero. Thus, with probability one, the realization of the network of techniques Φ is such that the set of entrepreneurs with access to two techniques that

[25] Chain feasibility rules out solutions to (8) in which $\lambda_{s(\phi)} = \lambda_{b(\phi)} = 0$. While there is always such a solution to (8), there is no feasible allocation of labor that would deliver the supplier's good at a marginal cost of zero.

[26] The restrictions placed by pairwise stability on entrepreneurs efficiencies are relatively weak. One clear example: For any Φ, there is always a pairwise stable equilibrium in which no entrepreneur produces, with $q_j = 0, \forall j$. Other examples can be constructed for special cases of Φ. Suppose that Φ is such that sets indexed by $n \in \{0, 1, ...\}$ and that any technique for which the buyer is in set n has a supplier in set $n+1$. Suppose that there is a pairwise-stable equilibrium in which entrepreneurs' efficiencies are $\{q_j\}$. Then for any $t \leq 1$ there is a pairwise stable equilibrium in which entrepreneurs' efficiencies are $\{\tilde{q}_j\}$, where $\tilde{q}_j = t^{\alpha^{-n_j}} q_j$, where j is in set n_j. Such equilibria are robust to deviations by any finite coalition but are not countably stable.

[27] Markups on sales to the household distort the consumption-leisure margin, but since labor is supplied inelastically there is no impact on the allocation of goods and labor. If labor were supplied elastically or if demand elasticities varied across goods, efficiency would require correcting the monopoly markups on sales to the household.

deliver the same efficiency has measure zero. Finally, payments between buyers and suppliers have no impact on the equilibrium allocation because all entrepreneurs remit profit to the household.[28]

2.5 Stability and Payoffs

Pairwise stability is sufficient to ensure that for every technique, $\frac{wl(\phi)}{1-\alpha} = \frac{\lambda_{s(\phi)}x(\phi)}{\alpha} = \lambda_{b(\phi)}y(\phi)$. Consequently, j's payoff can be decomposed into three parts, profit from sales to the household plus surplus from any buyers less surplus paid to suppliers:

$$\pi_j = \pi_j^0 + \sum_{\phi \in D_j} \tau(\phi) - \sum_{\phi \in U_j} \tau(\phi) \tag{9}$$

where $\pi_j^0 \equiv (p_j^0 - \lambda_j)y_j^0$ and $\tau(\phi) \equiv T(\phi) - \lambda_{s(\phi)}x(\phi)$. Given the arrangement and individual choices, π_j^0 is the profit from sales of good j to the household when good j is valued at j's marginal cost. $\tau(\phi)$ is the value of the payment above the cost of the intermediate inputs when those inputs are valued at the supplier's marginal cost.

While countable stability pins down the quantities traded, there is a dimension of multiplicity. For each pair there is a range of payments that are consistent with stability; there are many ways buyer-supplier pairs can split surplus.[29] Proposition 2 describes essential upper and lower bounds to the payments among entrepreneurs. These bounds apply to entrepreneurs whose supply chains do not cycle and do not overlap. Formally, we say entrepreneur j has an acyclic network if, for any other entrepreneur $\tilde{j}$, there is at most one undirected path of techniques linking j and $\tilde{j}$. This means that no supply chain that goes through j contains a cycle and that no two supply chains that pass through j cross paths at any other entrepreneur. Let J^* be the set of entrepreneurs with acyclic networks. With probability 1, Φ is such that J^* has measure 1.[30] Let Φ^* be the set of techniques for which

[28] If entrepreneurs made entry/exit decisions or chose the intensity with which to search for new techniques, the ex-post distribution of profit across entrepreneurs would play a more central role.

[29] The range is analogous to a Nash-bargaining set, with one important difference. Here, the split between a buyer and supplier determines how much surplus to there is to be split between the supplier and her supplier.

[30] For any entrepreneur j, the set of entrepreneurs $\{j'\}$ for which there is path from j to j' is countable and thus has measure zero. Therefore the probability that j has two paths to any single entrepreneur is also zero.

the buyer and supplier are members of J^*.

We define the surplus delivered by a technique to be the contribution of that technique to the profit of all other entrepreneurs. Let $\mathcal{B}(\phi)$ be the set of all entrepreneurs with supply chains that pass through ϕ. The surplus delivered by the technique ϕ is then

$$\mathcal{S}(\phi) = \sum_{j \in \mathcal{B}(\phi)} \pi_j^0 - \pi_{j\backslash\phi}^0$$

where $\pi_{j\backslash\phi}^0$ is the profit j would earn from sales to the household if it were unable to use any supply chain that passed through the technique ϕ.[31] Proposition 2 says that, for almost all techniques, countable stability requires that the payment of surplus is non-negative and no greater than the surplus of the technique.

Proposition 2. *In any countably-stable equilibrium, $\tau(\phi) \in [0, \mathcal{S}(\phi)]$ for each $\phi \in \Phi^*$. For any $\beta \in [0,1]$, there exists a countably-stable equilibrium in which $\tau(\phi) = \beta\mathcal{S}(\phi), \forall \phi \in \Phi^*$ and $\tau(\phi) = 0, \forall \phi \notin \Phi^*$.*

If there is a technique for which $\tau(\phi) < 0$, the supplier would deviate by dropping the contract, and the entire production chain would reduce production efficiently so that the supplier's marginal cost would remain unchanged. Similarly, if there is a technique for which $\tau(\phi) > \mathcal{S}(\phi)$, the buyer would drop the contract and purchase goods from the supplier associated with its next best technique. That entire supply chain would increase production efficiently so that that supplier's marginal cost is unchanged. For any entrepreneur downstream from the buyer, production would adjust efficiently in their supply chains so that they make efficient use of their best supply chain that does not pass through the technique dropped by the buyer.

The second part of Proposition 2 describes a class of countably-stable equilibria for which there is a natural notion of bargaining power. For each β, there is an equilibrium in which the fixed part of the payment of almost every technique is equal to a fraction β of the surplus

[31] Formally, define $\mathfrak{C}_{j\backslash\phi}$ be the set of chains to produce good j that do not pass through the technique ϕ. Define $q_{j\backslash\phi} \equiv \sup_{\mathfrak{c} \in \mathfrak{C}_{j\backslash\phi}} \prod_{n=0}^{\infty} [z^n(\mathfrak{c})]^{\alpha^n}$ to be the efficiency delivered to j by its best supply chain that does not pass through technique ϕ. Finally, let $\pi_{j\backslash\phi}^0 \equiv \frac{1}{\varepsilon-1}(q_{j\backslash\phi}/Q)^{\varepsilon-1} wL$.

generated by that technique.[32]

(4) implies that the key to solving for aggregate output is to characterize entrepreneurs' efficiencies. Given the network of techniques, Φ, this can be done using (5). One approach is to look for a vector of efficiencies $\{q_j\}_{j\in[0,1]}$ that is a fixed point of (5).[33] However, with a continuum of entrepreneurs, this is neither computationally feasible nor would it be particularly illuminating. Section 3 describes an approach which uses (5) along with the probabilistic structure of the model to sharpen the characterization of the optimal allocation.

3 Using the Probabilistic Structure to Characterize the Equilibrium

The economy's production possibilities—the techniques available to the different entrepreneurs—is stochastic. Given the realization of Φ, the contracting arrangement and individual production choices determine each entrepreneur's efficiency. This section uses the law of large numbers to solve for the *distribution* of efficiencies that is likely to arise given this probabilistic structure. While any individual entrepreneur's efficiency varies across realizations of the economy, the cross-sectional distribution of efficiencies does not. This section shows that the CDF of this cross-sectional distribution is the unique solution to a fixed point problem.

Given Φ, let $F(q)$ be the fraction of entrepreneurs with efficiency no greater than q in equilibrium. This function is endogenous and will need to be solved for. The strategy exploits the fact that F describes the distribution of an entrepreneur's efficiency and that of each of its potential suppliers.

What is the probability that an entrepreneur has efficiency no greater than q? This depends on how many techniques she discovers and the efficiency each of those techniques might deliver. The number of upstream technique each entrepreneur has, $|U_j|$, follows a Poisson distribution fully described by its mean, M.

[32] For the properties of the equilibrium described in Section 4 such as the size distribution and the cross-sectional correlations, one can safely ignore sets of measure zero.

[33] Taking logs of both sides gives an operator T, where the jth element of $T\left(\{\log q_j\}_{j\in[0,1]}\right)$ is $\max_{\phi\in U_j}\{\log z(\phi)+\alpha\log q_{s(\phi)}\}$. T satisfies monotonicity and discounting. If the support of z were bounded above, the operator would be a contraction on the appropriately bounded function space. This approach to solving the model is useful for simulations with a finite set of entrepreneurs.

Two elements determine the cost-effectiveness of each technique: (i) its productivity, $z(\phi)$, drawn from an exogenous distribution H, and (ii) the efficiency of the supplier, $q_{s(\phi)}$. Since the identity of the supplier is drawn uniformly, the probability that the supplier's efficiency is no greater than q_s is $F(q_s)$.

Let $G(q)$ be the probability that the efficiency delivered by a single random technique is no greater than q. Recalling that the efficiency delivered by the technique is $z(\phi)q^{\alpha}_{s(\phi)}$, this is:

$$G(q) = \int_{z_0}^{\infty} F\left((q/z)^{1/\alpha}\right) dH(z) \tag{10}$$

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

Now, the probability that, given all of its techniques, an entrepreneur has efficiency no greater than q is:

$$\Pr\left(q_j \leq q\right) = \sum_{n=0}^{\infty} \underbrace{\frac{M^n e^{-M}}{n!}}_{\Pr(n \text{ techniques})} \underbrace{G(q)^n}_{\Pr(\text{All } n \text{ techniques are} \leq q)} = e^{-M[1-G(q)]}$$

To interpret this last expression, note that if $M\left[1 - G(q)\right]$ is the mean of a Poisson distribution (the arrival of techniques that deliver efficiency better than q), then $e^{-M[1-G(q)]}$ is the probability of no such techniques.

The law of large numbers implies that with, probability one, $\Pr\left(q_j \leq q\right) = F(q)$.[34] Using the expression for $G(q)$ from equation (10) gives a fixed point problem for the CDF of efficiency F:

$$F(q) = e^{-M\int_{z_0}^{\infty}\left[1-F\left((q/z)^{1/\alpha}\right)\right]dH(z)} \tag{11}$$

This recursive equation is the key to characterizing the equilibrium.

Consider the space $\bar{\mathcal{F}}$ of right-continuous, non-decreasing functions $f : \mathbb{R}^+ \mapsto [0, 1]$, and

[34] The proof of Proposition 3 uses such a law of large numbers for a continuum of random variables described by Uhlig (1996). To use this, one must verify that entrepreneurs' efficiencies are pairwise uncorrelated. That is not immediately obvious in this context, as it is possible that two entrepreneurs' 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 entrepreneurs but only a countable set of those are in any of a given entrepreneur's potential supply chains. Therefore, for any two entrepreneurs, the probability that their supply chains overlap is zero.

consider the operator T on this space defined as

$$Tf(q) \equiv e^{-M\int_{z_0}^{\infty}\left[1-f\left((q/z)^{1/\alpha}\right)\right]dH(z)}$$

Online Appendix B constructively defines a subset $\mathcal{F} \subset \bar{\mathcal{F}}$ that depends on M and H. Proposition 3 shows that the equilibrium allocation is the unique fixed point of T on $\mathcal{F}$. The proof, contained in Online Appendix B, uses Tarski's fixed point theorem to show existence, which also provides a numerical algorithm to solve for it.

Proposition 3. *There exists a unique fixed point of T on $\mathcal{F}$, F. With probability one, F is the CDF of the cross-sectional distribution of efficiencies of every countably-stable equilibrium, and aggregate productivity is $Q = \left(\int_0^{\infty} q^{\varepsilon-1}dF(q)\right)^{\frac{1}{\varepsilon-1}}$.*

The qualitative features of the economy depend on whether the average number of techniques, M, is greater or less than one. If $M \leq 1$, there are so few techniques that the probability that any individual entrepreneur has access to a supply chain is zero.[35] In the more interesting case in which M exceeds the critical value of 1, there are at least three fixed points of the operator T on $\bar{\mathcal{F}}$, only one of which corresponds to the equilibrium.[36] Because of the multiplicity of fixed points, it is important to restrict the fixed point problem to the

[35] As shown in Online Appendix B.4, the probability that an entrepreneur has no viable chains is the smallest root ρ of $\rho = e^{-M(1-\rho)}$. For $M \leq 1$, the smallest root is one, while for $M > 1$, it is strictly less than one. This uses a standard result from the theory of branching processes. If $M < 1$, T is a contraction, with the unique solution $f = 1$. A phase transition when the average number of links per node crosses 1 is a typical property of random graphs, a result associated with the Erdos-Renyi Theorem. Kelly (1997) gives such a phase transition an economic interpretation.

[36] There are two solutions in which $f(q)$ is constant for all q which stem from the fact that equation (11) is formulated recursively. The first is $f(q) = 1$ for all q, which corresponds to zero efficiency (infinite marginal cost) for all goods; if the marginal cost of every input is infinite, then the marginal cost of each output is infinite as well. This allocation is feasible but not countably stable. A second, $f(q) = \rho \in (0, 1)$, implies infinite efficiency for entrepreneurs with supply chains and zero efficiency for those without; if inputs have zero cost then output can be produced at zero cost. As discussed in footnote 25, this does not correspond to a feasible allocation.

Each of these correspond to a fixed point of (5). Multiple solutions to first order conditions is actually a typical feature of models with roundabout production in which a portion of output is used simultaneously as an input. If the same good enters a technological constraint as both an input and an output, the price of that good will be on both sides of a first order condition. Consequently, the first order condition will be satisfied if the price takes the value of zero or infinity. One can usually sidestep this issue by finding an alternative way to describe the production technology, i.e., solving for final output as a function of primary inputs. Much of the work in the proof of Proposition 3 is in finding and characterizing such an alternative description of production possibilities.

function space $\mathcal{F}$, a space for which there is a unique fixed point which corresponds to the equilibrium.

Rather than characterizing how entrepreneurs' efficiencies directly depend on the set of production possibilities Φ, Proposition 3 shows how the cross-sectional distribution of efficiencies depends on the three primitives that generate Φ: the distribution from which technique-specific productivities are drawn, H; the average number of techniques per entrepreneur, M; and α. Section 4 uses this cross-sectional distribution to characterize how these primitives impact productivity and the organization of production.

4 Cross-Sectional Implications

This section studies how selection of suppliers shapes cross-sectional features of the economy such as how input-output links are distributed across entrepreneurs, the size distribution, and aggregate productivity.

To illustrate these implications, I focus on a parametric assumption that proves to be analytically tractable, allowing for closed-form expressions for the distribution of efficiencies and for aggregate output, and providing a transparent connection between the features of the environment and economic outcomes.

Assumption 2. *For each producer, the number of upstream techniques with match-specific productivity greater than z follows a Poisson distribution with mean $mz^{-\zeta}$, with $\zeta > \varepsilon - 1$.*

An economy that satisfies Assumption 2 can be interpreted as the limit of a sequence of economies that satisfies Assumption 1.[37] Consider an economy that satisfies Assumption 1 in which each technique's match-specific productivity is drawn from a Pareto distribution, $H(z) = 1 - (z/z_0)^{-\zeta}$. The shape parameter ζ governs the thickness of the right tail—smaller ζ corresponds to a thicker tail—and there is a minimum cutoff z_0. The sequence of economies takes $z_0 \to 0$, holding fixed the arrival rate of techniques above any productivity level, $M[1 - H(z)]$. Formally, define m to satisfy $M = mz_0^{-\zeta}$. m, a normalized measure of

[37]This type of distributional assumption can be traced back to at least Houthakker (1955), although he does not make a formal argument about a limit. Kortum (1997) shows that his model is well behaved asymptotically under this type of functional form assumption.

techniques, is defined this way so that for any $z \geq z_0$, the arrival rate of techniques above z is $M[1 - H(z)] = mz^{-\zeta}$. An economy that satisfies Assumption 2 is the limit of a sequence of such economies as $z_0 \to 0$ holding m fixed.[38]

Under Assumption 2, every solution F to equation (11) is the CDF of a Frechet distribution. To see this, note that $H(z) = 1 - \left(\frac{z}{z_0}\right)^{-\zeta}$ and $M = mz_0^{-\zeta}$ imply equation (10) becomes

$$\begin{aligned} M\left[1 - G(q)\right] &= mz_0^{-\zeta} \int_{z_0}^{\infty} \left[1 - F\left((q/z)^{1/\alpha}\right)\right] \zeta z_0^{\zeta} z^{-\zeta-1} dz \\ &= q^{-\zeta} m \int_0^{(q/z_0)^{1/\alpha}} \left[1 - F(x)\right] \alpha\zeta x^{\alpha\zeta-1} dx \end{aligned}$$

where the second line uses the change of variables $x = (q/z)^{1/\alpha}$. For any q, as $z_0 \to 0$, this expression approaches $q^{-\zeta}$ multiplied by a constant. Label this constant θ, so that equation (11) can be written as

$$F(q) = e^{-\theta q^{-\zeta}}$$

the CDF of a Frechet random variable. The mean of the distribution is increasing in θ, the location parameter, which was defined to satisfy $\theta = m \int_0^{\infty} \left[1 - F(x)\right] \alpha\zeta x^{\alpha\zeta-1} dx$. Using $F(q) = e^{-\theta q^{-\zeta}}$ and integrating gives $\theta = \Gamma(1-\alpha) m \theta^{\alpha}$, or more simply

$$\theta = \left[\Gamma(1-\alpha) m\right]^{\frac{1}{1-\alpha}} \tag{12}$$

where $\Gamma(\cdot)$ is the gamma function.[39]

The efficiency distribution inherits the tail behavior of the match-specific productivity draws; in the expression for F, the exponent ζ is the same as that of the Pareto distribution H. As one might expect, θ is increasing in the (normalized) number of techniques per entrepreneur: an entrepreneur with more options to choose from will tend to have higher efficiency. The other determinant of θ is α; its role will be discussed in detail below.

[38] Note that in this limit the arrival rate of techniques grows without bound ($M \to \infty$) while the distribution of match-specific productivity for any single technique deteriorates ($H(z) \to 1$). The normalization ensures that these occur at the same rate so that the limiting economy is well-behaved. In the limit, the measure of entrepreneurs with no techniques (e^{-M}) shrinks to zero.

[39] $\theta = \Gamma(1-\alpha) m \theta^{\alpha}$ has three non-negative roots: the one in equation (12), zero, and infinity. The latter two correspond to the two constant fixed points of equation (11), as discussed in footnote 36.

4.1 Stars and Superstars

An entrepreneur's efficiency depends on the arrival of upstream techniques and the availability of efficient supply chains. An entrepreneur's downstream techniques determine the demand for the entrepreneur's good for use as an intermediate input. This section studies how the interaction of upstream and downstream techniques determine an entrepreneur's size and contribution to aggregate productivity.

Proposition 4 characterizes how customers are distributed across suppliers.

Proposition 4. *Suppose Assumption 2 holds.*

1. *Among entrepreneurs with efficiency q, the number of actual customers follows a Poisson distribution with mean $\frac{m}{\theta}q^{\alpha\zeta}$.*

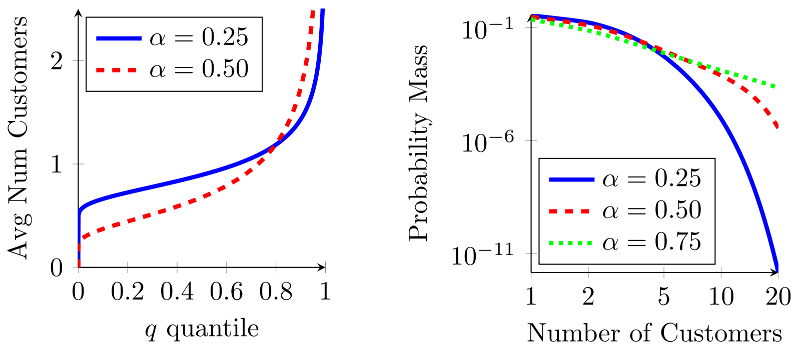
2. *Among all entrepreneurs, the distribution of customers asymptotically follows a power law with exponent $1/\alpha$: $\Pr(\#\ \textit{customers} \geq n) \sim \frac{1}{\Gamma(1-\alpha)^{1/\alpha}} n^{-1/\alpha}$.*

Proof Sketch. For any technique, let $\tilde{F}(x)$ be the probability that the potential buyer has no *other* techniques that deliver efficiency better than x. Since all such potential buyers have at least one upstream technique, this is $\tilde{F}(x) \equiv \frac{\sum_{n=1}^{\infty} G(x)^{n-1} \frac{e^{-M}M^n}{n!}}{1-e^{-M}} = \frac{F(x)-e^{-M}}{G(x)(1-e^{-M})}$. Note that $\lim_{z_0 \to 0} \tilde{F}(q) = F(q)$.

Consider an entrepreneur with efficiency q. A single downstream technique with productivity z would deliver efficiency zq^α to its potential customer, and $\tilde{F}(zq^\alpha)$ is the probability that the potential customer has no better alternative techniques. Since the number of potential customers is Poisson with mean M, the number of actual customers follows a Poisson distribution with mean $M \int_0^\infty \tilde{F}(zq^\alpha) dH(z)$. Under Assumption 2 this equals $m \int_0^\infty F\left(zq^\alpha\right) \zeta z^{-\zeta-1} dz$, and integrating yields the first result.

Integrating over entrepreneurs of different efficiencies, the mass of entrepreneurs with n customers is $\int_0^\infty \frac{e^{\frac{m}{\theta}q^{\alpha\zeta}}\left(\frac{m}{\theta}q^{\alpha\zeta}\right)^n}{n!} dF(q) = \int_0^\infty \frac{1}{n!} e^{-\left(\frac{w^{-\alpha}}{\Gamma(1-\alpha)}\right)} \left(\frac{w^{-\alpha}}{\Gamma(1-\alpha)}\right)^n e^{-w} dw$. The second result uses this along with a known result about tail indices of Poisson mixtures. ■

These properties are illustrated in Figure 2. Figure 2a shows the average number of customers at each quantile in the efficiency distribution for different values of α. As one



(a) Conditional Means (b) Unconditional Distribution

Figure 2: Distribution of Customers

Figure 2a shows the mean number of actual customers for each quantile in the efficiency distribution. Figure 2b gives the mass of entrepreneurs with n customers on a log-log plot. Under Assumption 2, the curves in each plot depend only on α.

would expect, each curve in Figure 2a is increasing, indicating high-efficiency entrepreneurs attract more customers.

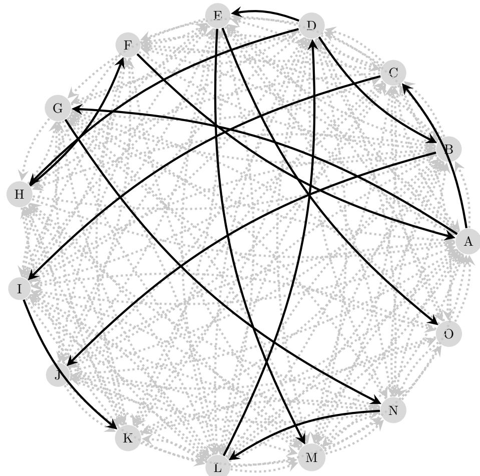
If α is larger, the high-efficiency entrepreneurs capture an even larger share of customers. There is a single crossing property that is evident in Figure 2a: with higher α, the expected number of customers is more steeply increasing with efficiency.[40]

Why do higher efficiency entrepreneurs attract proportionally more customers when α is higher? Recall that the efficiency delivered by a single technique is $z(\phi)q^{\alpha}_{s(\phi)}$. A technique is more likely to be used by the potential customer when the technique has high match-specific productivity and when the supplier has high efficiency. α is thus the elasticity of the efficiency delivered by a technique to the supplier's efficiency. When α is low, the supplier's efficiency is less important relative to the match-specific productivity because the impact on the technique's efficiency is muted. In contrast, when α is large, higher efficiency producers are more likely to be selected by each of their potential customers.[41]

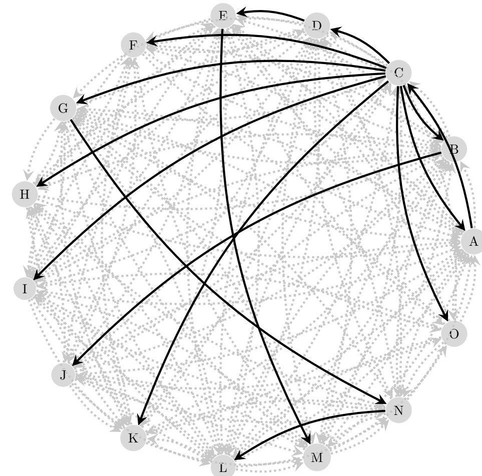
The second result of Proposition 4 describes the unconditional distribution of customers across entrepreneurs, and this is plotted in Figure 2b for different values of α. When α

[40] Among entrepreneurs at the ϱ-th quantile of the efficiency distribution (those with efficiency $F^{-1}(\varrho)$), the distribution of customers is Poisson with mean $\frac{\log[1/\varrho]^{-\alpha}}{\Gamma(1-\alpha)}$. Taking the log of this expression, it is easy to see that, across α, this satisfies a single crossing property; the curve is increasing more sharply for higher α.

[41] Another way to see this is to study how the entrepreneurs' efficiency covary with number of customers. Online Appendix C.1 shows that $\frac{\text{Cov}(\log q, \#\text{ customers})}{\text{St. Dev.}(\log q)} = \frac{\sqrt{6}}{\pi}\int_0^1 \frac{x^{-\alpha}-1}{1-x}dx$ which is increasing in α.



(a) $\alpha = 0.1$



(b) $\alpha = 0.9$

Figure 3: Equilibrium Supply Chains and α

Figure 3 shows entrepreneurs' choices of techniques. The set of techniques, Φ, is held fixed; the only difference is the value of α. The dark edges represent techniques that are used. $M = 15$ and $H(z) = 1 - z^{-2}$ for $z \geq 1$.

is larger, the distribution has a thicker tail, as the high-efficiency entrepreneurs are more likely to attract a disproportionate share of their potential customers. In other words, when α is large, the equilibrium network features more star suppliers—entrepreneurs with many customers. This is illustrated more concretely by Figure 3. In each figure, the set of techniques, Φ, is exactly the same, but α is larger in panel (b).

The model thus provides a mechanism for a skewed cross-sectional distribution of links which compliments that of the influential preferential attachment model of Barabasi and Albert (1999). Here, the extra variation stems from selection. Because entrepreneurs differ in marginal cost, they also differ in their likelihood of being selected by potential customers. Some of those entrepreneurs with many potential customers are able to offer lower prices and win over a larger fraction of those potential customers. While the distribution of *potential* customers follows a Poisson distribution, the distribution of *actual* customers exhibits more variation and asymptotically follows a power law.

Both the conditional means and unconditional distribution of customers depend on a single parameter, α. Notably absent is the parameter ζ which governs variation in entrepreneurs' efficiencies: the standard deviation of $\log q$ is $\frac{1}{\zeta}\frac{\pi}{\sqrt{6}}$. In particular, even if there is arbitrarily little variation in entrepreneurs' marginal costs (i.e., ζ is large) the distribution of customers across entrepreneurs would follow a power law with exponent $1/\alpha$, and entrepreneurs with high efficiency would have a disproportionate share of customers.

Why is the distribution of customers independent of ζ? ζ actually plays two offsetting roles. The distribution of customers depends on both the variation in cost of inputs and the likelihood that a potential customer would forgo a technique with better match-specific productivity for one that relies on less expensive inputs. The latter depends on magnitude of variation in technique-specific productivity draws. If all techniques had the same productivity, each potential buyer would simply select the technique with the least expensive input, and the distribution of customers would be heavily skewed toward the most efficient entrepreneurs. In contrast, if all inputs could be purchased at the same cost, each buyer would simply choose the technique with the best fit (highest z). In that case customers would be relatively evenly distributed across entrepreneurs; the customer distribution would be Poisson. Thus a key driver of the heterogeneity in the distribution of customers is the magnitude of variation in match-specific productivity draws relative to the magnitude variation in suppliers' marginal costs.

A noteworthy feature of the roundabout nature of production is that the parameter ζ drives variation in both. When ζ is smaller so that there is more variation in productivity draws, there is endogenously more variation in suppliers' cost; the suppliers' costs depend on the match-specific productivity draws of their own upstream techniques.[42] Thus a smaller ζ compresses the distribution of customers by increasing the variation in match-specific productivity draws but also expands the distribution of customers by increasing the variation of potential suppliers' cost of production. These two channels exactly offset.

The Role of Functional Forms

At this point it is worth commenting on the role of Assumption 2. One reason to focus on the Assumption 2 is that it is a special case in which the model has determinate cross-sectional patterns that do not depend on the scale of the economy. For example the distribution of

[42]As mentioned above, the distribution of efficiencies inherits the tail of the distribution of productivity draws.

customers is independent of m. If m were larger, each entrepreneur would have more potential customers. However, on average, the number of customers does not increase because of a countervailing force: a larger m also raises the number of alternatives available to each of those potential customers. These offset each other, and the scale-free nature of Assumption 2 ensures that they offset exactly.[43]

However, these assumptions are not crucial to the main results. Two results can be shown. The first result is fairly intuitive: because customers base their selection of suppliers partly on prices, the distribution of actual customers exhibits more variation than the distribution of potential customers.[44] In other words, selection on prices is a mechanism that leads to increased heterogeneity in engagement in input-output linkages.

The second result is that the prevalence of star suppliers—and the entire distribution of customers—is independent of the variation in entrepreneurs' marginal cost. To see that this result does not depend on Assumption 2, consider the thought experiment of an economy indexed by γ. Relative to the benchmark, in the γ-economy each z is replaced by $\hat{z} \equiv z^{\gamma}$, so that $H_{\gamma}(\hat{z}) \equiv H\left(\hat{z}^{1/\gamma}\right)$; the benchmark economy is recovered by setting $\gamma = 1$. In the γ-economy, the efficiency delivered by supply chain $\mathfrak{c}$ is $\prod_{n=0}^{\infty}[\hat{z}^{n}(\mathfrak{c})]^{\alpha^{n}} = \prod_{n=0}^{\infty}[z^{n}(\mathfrak{c})^{\gamma}]^{\alpha^{n}} = \left(\prod_{n=0}^{\infty}[z^{n}(\mathfrak{c})]^{\alpha^{n}}\right)^{\gamma}$. There are two implications. First, $\hat{q}_j = q_j^{\gamma}$, which means that the dispersion of the log of entrepreneurs' efficiencies is proportional to γ. Second, the ordinal ranking of the supply chains is independent of γ. As a consequence the economy's equilibrium input-output linkages are independent of γ.

4.2 The Distribution of Employment

The distribution of employment across producers has long generated considerable interest from economists.[45] This section builds on Section 4.1 to characterize the determinants of the

[43] Recall that Assumption 2 can be interpreted as the limit of a sequence of economies as $z_0 \to 0$. Online Appendix C.1 describes how the distribution of customers changes as this sequence converges to its limit.

[44] Formally the coefficient of dispersion (the ratio of variance to mean) of actual customers is larger than that of potential customers. This follows from the fact that the variance of a Poisson distribution is the same as its mean and the law of total variances. For potential customers, the mean and variance of potential customers is M, so the coefficient is 1. Let n_j be the number of j's actual customers. $\mathbb{E}[n] = 1$, and $\text{Var}(n) = \text{Var}(\mathbb{E}[n|q]) + \mathbb{E}[\text{Var}(n|q)]$. The first term is weakly positive, and the second term is $\mathbb{E}[\text{Var}(n|q)] = \mathbb{E}[\mathbb{E}(n|q)] = 1$.

[45] Among the many, see Lucas Jr (1978), Jovanovic (1982) Hopenhayn (1992), Axtell (2001), Klette and Kortum (2004), Luttmer (2007), Rossi-Hansberg and Wright (2007), and Arkolakis (2011).

distribution of employment across entrepreneurs. An entrepreneur produces output to sell to the household for consumption and to other entrepreneurs for intermediate use. Entrepreneur j's choice of employment ultimately depends on the indirect demand for j's good by all other entrepreneurs that produce using supply chains that go through j.

Because an entrepreneur's size depends on the size of each of its customers, the distribution of employment can be characterized recursively. This characterization relies on the fact that a customer's size is sufficient to summarize the indirect demand for j's labor by any other entrepreneurs downstream from that customer. Let $\mathcal{L}(\cdot)$ be CDF of the overall size distribution and let $\mathcal{L}(\cdot|q)$ be the CDF of the conditional size distribution among entrepreneurs with efficiency q, so that $\mathcal{L}(l) = \int_0^\infty \mathcal{L}(l|q)dF(q)$. Proposition 5 describes, for most parameter values, the model's implications for these distributions.

Proposition 5. *Suppose that Assumption 2 holds and that* $\rho \equiv \min\left\{\frac{1}{\alpha}, \frac{1}{(\varepsilon-1)/\zeta}\right\}$ *is not an integer. Then right tails of the overall and conditional distributions of employment follow power laws with exponent* ρ*:*

1. $1 - \mathcal{L}(l) \sim Kl^{-\rho}$

2. $1 - \mathcal{L}(l|q) \sim K\frac{mq^{\alpha\zeta}}{\theta}(\alpha l)^{-\rho}$

where $K \equiv \frac{1}{1-\alpha^\rho}\frac{\left((1-\alpha)\mathbb{I}_{\frac{\varepsilon-1}{\zeta}\geq\alpha}+\alpha\mathbb{I}_{\alpha\geq\frac{\varepsilon-1}{\zeta}}\right)^\rho}{\Gamma(1-\rho^{-1})^\rho}$.

Proof Sketch. Because employment is used to produce output for each destination—for each customer and for the household—rather than working with the overall and conditional size distributions, $\mathcal{L}$ and $\{\mathcal{L}(\cdot|q)\}$, it will be easier to work with their respective Laplace-Stieltjes transforms, $\hat{\mathcal{L}}(s) \equiv \int_0^\infty e^{-sl}d\mathcal{L}(l)$ and $\hat{\mathcal{L}}(s|q) \equiv \int_0^\infty e^{-sl}d\mathcal{L}(l|q)$. The first part of the proof derives fixed point problems for these transforms. Under Assumption 1, the transform of the conditional employment distribution among those with efficiency q is

$$\begin{aligned}\hat{\mathcal{L}}(s|q) &= e^{-s(1-\alpha)(q/Q)^{\varepsilon-1}L}\sum_{n=0}^{\infty}\frac{e^{-M}M^n}{n!}\left[\int_0^\infty \hat{\mathcal{L}}(\alpha s|zq^\alpha)\tilde{F}(zq^\alpha)dH(z)\right]^n \\ &= e^{-s(1-\alpha)(q/Q)^{\varepsilon-1}L}e^{-M\int_0^\infty\left[1-\hat{\mathcal{L}}(\alpha s|zq^\alpha)\right]\tilde{F}(zq^\alpha)dH(z)}\end{aligned}$$

where $e^{-s(1-\alpha)(q/Q)^{\varepsilon-1}L}$ is the transform of labor used for the household, $\hat{\mathcal{L}}(\alpha s|zq^{\alpha})$ is the transform of labor used for a potential customer with match-specific productivity z, and $\tilde{F}(zq^{\alpha})$ is the probability that such a potential customer has no better alternatives, as described in the proof of Proposition 4. Imposing Assumption 2 yields $\hat{\mathcal{L}}(s|q) = e^{-s(1-\alpha)(q/Q)^{\varepsilon-1}L}e^{-\frac{m}{\theta}q^{\alpha\zeta}[1-\hat{\mathcal{L}}(\alpha s)]}$ Integrating over entrepreneurs with different efficiencies, and using the change of variables $t = \theta q^{-\zeta}$ gives a single fixed point problem for the overall distribution of employment:

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

The second part of the proof uses a tauberian theorem to relate the slope of the transform at zero to the slope of the right tail of the employment distribution.[46] ∎

Proposition 5 describes the determinants of the employment distribution, and the proof shows that its shape depends only on α and $\frac{\varepsilon-1}{\zeta}$. α matters because it determines the distribution of customers and because it determines how much labor must be used to produce intermediate inputs for a customer. $\frac{\varepsilon-1}{\zeta}$ is a composite of two parameters, the elasticity of substitution across varieties in consumption and ζ, the shape parameter of the distribution of efficiencies. In combination, these parameters determine the shape of the distribution of sales to the household. When ζ is small, there is more dispersion in prices that the household faces, and when ε is high, the household is more willing to substitute toward goods with lower prices.

The first result of Proposition 5 describes the determinants of the right tail of the size distribution. The slope of the right tail is governed by either labor used to make intermediate inputs or by labor used to make final consumption. Across entrepreneurs, the upper tail of labor used to make intermediate inputs for others has a Pareto tail with exponent $1/\alpha$, while labor used to make consumption has a Pareto tail of $\zeta/(\varepsilon-1)$. The proposition says that one of these features will dominate and determine the upper tail of the overall size distribution.

One way to understand this is to note that the conditional transform implies that among entrepreneurs with efficiency q, the average employment is $(1-\alpha)(q/Q)^{\varepsilon-1} + \alpha\frac{m}{\theta}q^{\alpha\zeta}$.[47] The

[46] The particular tauberian theorem used is valid when ρ is not an integer. While I have not been able to prove it, a reasonable conjecture is that the theorem also holds when ρ is an integer.

[47] The n-th moment of a distribution is $(-1)^n$ times the n-th derivative of the transform.

first term is labor used to sell goods to the household, the second to customers (recall $\frac{m}{\theta}q^{\alpha\zeta}$ is the expected number of customers). When α is small, the sales to the household are more important. When α is larger, sales of intermediates becomes more important, but also these sales become more concentrated in star suppliers: $\frac{m}{\theta}q^{\alpha\zeta}$ becomes more steeply increasing in q. When $\alpha > \frac{\varepsilon-1}{\zeta}$, the increases in average employment across entrepreneurs with different efficiencies becomes dominated by the second term, the concentration of sales of intermediates.

The second result of Proposition 5 says that even among entrepreneurs with the same efficiency, the tail of the conditional distribution of employment follows a power law with the same slope as the overall employment distribution. In other words, there is a lot of variation in size even among entrepreneurs with the same marginal cost. Among these entrepreneurs, there is no variation in sales to the household and the distribution of customers follows a Poisson distribution, which has a thin tail. What then drives variation in employment? The granularity of customers: some customers are much larger than others. If, for example, an entrepreneur's customer is a star supplier, the entrepreneur will require a large mass of labor to make intermediate inputs for that customer. Under Assumption 2, the distribution of customer size has a power law tail. The possibility of having a customer at the upper end of the overall size distribution causes the conditional size distribution to inherit this power law tail.[48]

Put differently, suppose $\alpha > \frac{\varepsilon-1}{\zeta}$. The first result of the proposition implies that even when there is little variation in entrepreneurs' efficiencies (ζ large) there can be a lot of variation in size in the cross-section. To use the language of Acemoglu et al. (2012), this is driven by first-order interconnections: some entrepreneurs have many more customers than others. The second result further implies that there is a lot of variation in size even among entrepreneurs with the same efficiency. While there is little variation in first-order

[48] Another way to see that the feature that dominates the conditional distribution is the possibility of having a large customer is to compare the scales of the right tails of the overall and conditional employment distributions. Rather than l^ρ, the tail of the conditional size distribution is scaled by $(\alpha l)^\rho$; recall that if a customer uses l units of labor, its supplier uses αl units of labor in producing the intermediate inputs for that customer. In addition, the tail is scaled by $\frac{mq^{\alpha\zeta}}{\theta}$, the expected number of customers. This is closely related to the fact that if N independent random variables X_n are distributed so that $\Pr(X_n > x) \sim x^{-\rho}$ then $\Pr(\sum_{n=1}^{N} X_n > x) \sim N x^{-\rho}$.

interconnections across these entrepreneurs, second-order interconnections drive the variation in size: some of these entrepreneurs' customers have many customers.

Online Appendix D presents some preliminary evidence on the relationship between intermediate input shares and the right tails of size distributions. Using data on the size distribution among producers in France and the United States, I ask whether industries with higher intermediate input shares have size distributions with thicker tails. First, I use estimates from Di Giovanni et al. (2011) of tail exponents of the distribution of revenue among firms for each industry in France. In line with the theory, those industries with higher intermediate input shares tend to have thicker tails. Second, I estimate tail exponents for the distribution of employment among establishments and among firms in the United States using data extracts reported by Rossi-Hansberg and Wright (2007). While the point estimates suggest that industries with higher intermediate input shares have employment distributions with thicker tails, the estimates are not precise enough to distinguish statistically from zero.

4.3 Matching Patterns

In environments with matching, one frequently studied property is whether matches tend to exhibit positive or negative assortative matching. Proposition 6 shows that an implication of the model is that whether buyers and suppliers match assortatively depends on which attribute one compares.

Proposition 6. *Under Assumption 2, among techniques that are used in equilibrium:*

1. $\text{Cov}\left(\log q_{b(\phi)}, \log q_{s(\phi)}\right) = 0$
2. $\text{Cov}\left(n_{b(\phi)}, n_{s(\phi)}\right) = 0$, *where n_j is the number of j's customers*
3. $\text{Cov}\left(l_{b(\phi)}, l_{s(\phi)}\right) > 0$

Proof. We first show that the efficiency of a customer is independent of the efficiency of the supplier. If an entrepreneur with efficiency q_s has a downstream technique with productivity z, the potential buyer selects that technique with probability $\tilde{F}(zq_s^{\alpha})$ (defined in Proposition 4). Since the average number of downstream techniques is M, the fraction of actual customers whose efficiency is no greater than q is

$$\Pr(q_b < q | q_s, \text{technique is used}) = \frac{M \int_{z_0}^{q/q_s^\alpha} \tilde{F}(z q_s^\alpha) dH(z)}{M \int_{z_0}^{\infty} \tilde{F}(z q_s^\alpha) dH(z)}$$

Imposing Assumption 2 and noting that under this assumption $\tilde{F}(q) = F(q) = e^{-\theta q^{-\zeta}}$

$$\Pr(q_b < q | q_s, \text{technique is used}) = \frac{\int_0^{q/q_s^\alpha} e^{-\theta (z q_s^\alpha)^{-\zeta}} \zeta z^{-\zeta-1} dz}{\int_0^{\infty} e^{-\theta (z q_s^\alpha)^{-\zeta}} \zeta z^{-\zeta-1} dz} = e^{-\theta q^{-\zeta}}$$

which is independent of q_s. The second result follows because q_s and q_b are uncorrelated and the probability of being selected as a supplier depends only on efficiency. For the third result, if the buyer uses l_b units of labor, the supplier requires αl_b units of labor to make the intermediate inputs for that supplier. Since the other components of the supplier's labor are uncorrelated with l_b, the covariance of buyers' and suppliers' labor is $\text{Cov}(l_s, l_b) = \text{Cov}(\alpha l_b, l_b) = \alpha \text{Var}(l) > 0$.[49] ■

Across all techniques, those that use goods produced by higher-efficiency suppliers tend to deliver higher efficiency to the potential buyer. The first result of the proposition shows that, among techniques that are actually used in equilibrium, this correlation is quite different. The key mechanism is selection.

An entrepreneur selecting which of her techniques to use is unlikely to choose one that is associated with a low-efficiency supplier unless the match-specific productivity is unusually high. However, she may choose a technique associated with a high-efficiency supplier even if the match-specific productivity is low; the technique would still be relatively cost-effective because the intermediate inputs can be acquired at a low cost. Together, these imply that among techniques that are used in equilibrium, those associated with lower-efficiency suppliers tend to have higher match-specific productivity. A consequence is that among techniques that are used in equilibrium, suppliers' and buyers' efficiencies are uncorrelated.[50]

Second, the proposition shows that suppliers with more customers do not tend to have buyers with more customers. This is simply a corollary of the first result: the frequency with

[49](13) implies that the variance of employment is finite only if α and $\frac{\varepsilon-1}{\zeta}$ are small enough.

[50]The fact that the covariance is exactly zero depends on Assumption 2.

which an entrepreneur is selected as a supplier depends only on her efficiency. Since buyers' and suppliers' efficiencies are uncorrelated, there is no correlation in the rate at which they are selected as suppliers.

The third result indicates, however, that the sizes of buyers and suppliers, as measured by employment, are positively correlated in equilibrium. While the supplier's sales to the household and to other customers are uncorrelated with the size of the buyer, the supplier's labor used to produce intermediate inputs for the buyer is perfectly correlated with the buyer's size. If the buyer is large, the supplier will require a large amount of labor to satisfy that buyer's demand for intermediate inputs.

Nevertheless, these results together underscore that a note of caution is required when discussing whether matches exhibit positive assortative matching; two natural measures–correlation of buyers' and suppliers' marginal costs or of sizes–yield different answers.

4.4 The Cost Share of Intermediate Inputs

In most models with roundabout production, the cost share of intermediates is important because it corresponds to the input-output multiplier. This section shows that, in the environment studied here, the cost share is endogenous, and the next section shows that the cost share may not correspond to the input-output multiplier.

Each entrepreneur makes payments to its supplier and to labor. Entrepreneur j's cost share of intermediates is $\frac{\sum_{\phi \in U_j} T(\phi)}{\sum_{\phi \in U_j} [wl(\phi)+T(\phi)]}$. Similarly, the aggregate cost share of materials is $\frac{\int_0^1 \sum_{\phi \in U_j} T(\phi) dj}{\int_0^1 \sum_{\phi \in U_j} [wl(\phi)+T(\phi)] dj}$. Recall that a payment to a supplier can be decomposed into variable and fixed components, $T(\phi) = \lambda_{s(\phi)} x(\phi) + \tau(\phi)$, and that pairwise stability implies that the cost of labor used with a technique is proportional to the variable component of the payments to the supplier, $\lambda_{s(\phi)} x(\phi) = \frac{\alpha}{1-\alpha} wl(\phi)$. The aggregate cost share therefore equals

$$\frac{\int_0^1 \sum_{\phi \in U_j} \left[\frac{\alpha}{1-\alpha} wl(\phi) + \tau(\phi)\right] dj}{\int_0^1 \sum_{\phi \in U_j} \left[wl(\phi) + \frac{\alpha}{1-\alpha} wl(\phi) + \tau(\phi)\right] dj}$$

Rearranging and using the labor market clearing condition $\int_0^1 \sum_{\phi \in U_j} l(\phi)dj = L$ gives

$$\frac{\alpha wL + (1-\alpha)\int_0^1 \sum_{\phi \in U_j} \tau(\phi)dj}{wL + (1-\alpha)\int_0^1 \sum_{\phi \in U_j} \tau(\phi)dj}$$

One immediate result is that the intermediate input cost share is weakly larger than α, because in any countably stable equilibrium, $\tau(\phi)$ is almost surely weakly positive.[51] Proposition 7 characterizes the aggregate intermediate input cost share under Assumption 2 for the equilibrium in which suppliers receive a fraction β of surplus of each technique.

Proposition 7. *Under Assumption 2, in a countably-stable equilibrium in which $\tau(\phi) = \beta\mathcal{S}(\phi)$, the aggregate cost share of intermediates is*

$$\frac{\alpha + \beta/\zeta}{1 + \beta/\zeta} \tag{14}$$

Average revenue among those with efficiency q is $\left[\frac{\varepsilon}{\varepsilon-1}(q/Q)^{\varepsilon-1} + \frac{\alpha+\beta/\zeta}{1-\alpha}\frac{m}{\theta}q^{\alpha\zeta}\right]wL$.

Proof Sketch. Since techniques can be enumerated by either their buyers or suppliers, $\int_0^1 \sum_{\phi \in U_j} \tau(\phi)dj = \int_0^1 \sum_{\phi \in D_j} \tau(\phi)dj = \beta \int_0^1 \sum_{\phi \in D_j} \mathcal{S}(\phi)dj$. The rest of the proof shows that $\int_0^1 \sum_{\phi \in D_j} \mathcal{S}(\phi)dj = \frac{1}{1-\alpha}\frac{1}{\zeta}wL$.

Define $r_j \equiv \sum_{\phi \in D_j} \mathcal{S}(\phi)$ to be the total surplus of all of j's downstream techniques. Define $v_j \equiv \pi_j^0 + r_j$ to be the surplus of entrepreneur j; if j were unable to produce, v_j would be the total change in profit among j and all entrepreneurs downstream from j. Let $V(q)$ and $R(q)$ be the averages of v_j and r_j respectively among entrepreneurs with efficiency q. Among techniques downstream from such entrepreneurs with match-specific productivity z and for which the buyer's best alternative delivers efficiency $\tilde{q}$, the average surplus is $V(\max\{zq^\alpha, \tilde{q}\}) - V(\tilde{q})$. $R(q)$ is found by averaging over realizations of z, $\tilde{q}$, and the number of downstream techniques:

$$R(q) = M\int_{z_0}^{\infty}\int_0^{\infty} \left[V(\max\{zq^\alpha, \tilde{q}\}) - V(\tilde{q})\right] d\tilde{F}(\tilde{q})dH(z)$$

[51] While it is assumed that the labor market is competitive, collective bargaining could drive a wedge between the wage and the marginal cost of labor, pushing down the cost share of intermediate inputs.

where, as in the proof of Proposition 4, $\tilde{F}(\tilde{q})$ is the probability that the buyer has no alternative technique that delivers efficiency better than $\tilde{q}$. Imposing Assumption 2 (which also implies $\tilde{F}(\tilde{q}) \to F(\tilde{q})$) and making the change of variables $u = zq^\alpha$ gives

$$R(q) = m \int_0^\infty \int_0^\infty [V(\max\{zq^\alpha, \tilde{q}\}) - V(\tilde{q})] \, dF(\tilde{q}) \zeta z^{-\zeta-1} dz = \frac{m}{\theta} q^{\alpha\zeta} \rho$$

where $\rho \equiv \theta \int_0^\infty \int_0^\infty [V(\max\{u, \tilde{q}\}) - V(\tilde{q})] \, dF(\tilde{q}) \zeta u^{-\zeta-1} du$. V can thus be expressed as $V(q) = \frac{1}{\varepsilon-1}(q/Q)^{\varepsilon-1} wL + \frac{m}{\theta} q^{\alpha\zeta}\rho$. Substituting this into the definition of ρ and solving for ρ yields $\rho = \frac{1}{1-\alpha}\frac{1}{\zeta} wL$. Finally, we have that $\int_0^1 \sum_{\phi \in D_j} \mathcal{S}(\phi) dj = \int_0^\infty R(q) dF(q) = \rho$.

Revenue of entrepreneur j is $\pi_j^0 + \beta r_j + \frac{wl_j}{1-\alpha}$. Average revenue of those with efficiency q is thus $\frac{1}{\varepsilon-1}(q/Q)^{\varepsilon-1} wL + \beta R(q) + \frac{\int_0^\infty l d\mathcal{L}(l|q)}{1-\alpha}$. Using the expression for $R(q)$ from above and for $\int_0^\infty l d\mathcal{L}(l|q)$ from Section 4.2 gives the result. ■

If buyers have all of the bargaining power ($\beta = 0$) then the cost share is simply α. In that case the cost per unit of the intermediate input equals the buyer's (and supplier's) shadow value. If supplier's have more bargaining power, the cost share of intermediate inputs rises. However, the payment to a supplier is limited by the buyer's next best option. Under Assumption 2, $\frac{1}{\zeta}$ is a measure of the variation in match-specific productivities. When ζ is smaller, there is typically a larger distance between the efficiency delivered by each buyer's best and second-best techniques. This raises the surplus to be split between the buyer and the supplier, and increases the buyer's cost share of intermediates.

4.5 Aggregate Output

Aggregate output in the economy is $Y^0 = QL$. With the functional forms, aggregate productivity is $Q = \left(\int_0^\infty q^{\varepsilon-1} dF(q)\right)^{\frac{1}{\varepsilon-1}} = \theta^{1/\zeta}\Gamma\left(1 - \frac{\varepsilon-1}{\zeta}\right)^{\frac{1}{\varepsilon-1}}$. Combining this with the expression for θ from (12) gives an expression for the household's consumption:

$$Y^0 = \left[\Gamma\left(1 - \frac{\varepsilon-1}{\zeta}\right)^{\frac{\zeta}{\varepsilon-1}} \Gamma(1-\alpha)^{\frac{1}{1-\alpha}} m^{\frac{1}{1-\alpha}}\right]^{\frac{1}{\zeta}} L \tag{15}$$

There are several immediate implications. First, as in Kortum (1997), aggregate output increases with more techniques, m.[52] In an economy with more techniques, entrepreneurs tend to have larger sets of supply chains to choose from and, hence, are more likely to have found an efficient one. When ζ is small, the distribution of productivity draws has a thicker upper tail. The exponent $\frac{1}{\zeta}$ scales (exponentially) the productivity of any technique, and consequently each entrepreneur's efficiency. The term $\Gamma\left(1-\frac{\varepsilon-1}{\zeta}\right)^{\frac{\zeta}{\varepsilon-1}}$ reflects the household's ability to consume more of less expensive goods. ε, the elasticity of substitution across varieties, measures the household's willingness to substitute toward lower cost goods, while ζ indicates how much cheaper these lower cost goods are.

Second, the exponent $\frac{1}{1-\alpha}$ that appears in several places is an input-output multiplier that shows up in any model with roundabout production. α determines the extent to which lower input prices feed back into lower cost of production. A notable difference between this and the standard input-output multiplier is that in this context α not necessarily the same as the cost share of intermediate inputs; if suppliers have some bargaining power ($\beta > 0$) then α is smaller than the cost share.

A separate, more interesting, role is that α determines the composition of the entrepreneurs supplying intermediate inputs. Recall from Section 4.1 that α determines the frequency with which the lower-cost producers are selected as suppliers. In other words, when α is closer to one, the lowest cost producers are more likely to become star suppliers and be more relevant for aggregate production. At a more fundamental level, the supply chains used to produce each good are more likely to be routed through the higher-productivity techniques. Aggregate output is higher because these higher-efficiency techniques are used more intensively. Mathematically, this shows up in the term $\Gamma(1-\alpha)$.[53]

In sum, when α is high, each supplier is able to pass through cost savings to its customers at a higher rate *and* supply chains used in equilibrium are more likely to be routed through higher efficiency techniques.

[52]In the special case in which $\alpha = 0$, input-output relationships play no role and the expression for aggregate output is the same as Kortum (1997).

[53]Note that $\Gamma(x)$ is decreasing on $(0, 1)$, with $\lim_{x\searrow 0}\Gamma(x) = \infty$ and $\Gamma(1) = 1$. To get a sense of the magnitude, a change in α from 0.55 to 0.65 increases the standard multiplier $\frac{1}{1-\alpha}$ from 2.2 to 2.9, increases $\Gamma(1-\alpha)$ by a factor of 1.3, and increases $\Gamma(1-\alpha)^{\frac{1}{1-\alpha}}$ by a factor of 3.2.

5 Conclusion

This paper developed a theory of the formation of an economy's input-output architecture and characterized the implications for the size distribution, organization of production, and productivity. When intermediate goods are more important in production, activity becomes more concentrated in star suppliers, even when differences across producers in marginal cost are arbitrarily small. This raises aggregate productivity as supply chains are more likely to be routed through higher productivity techniques, and also increases the market concentration in sales of intermediate goods. While research documenting patterns of micro linkages is at an early stage, the model provides both testable implications and an organizing framework to guide empirical work as new datasets emerge.

That being said, there are a number of ways in which the model could be enriched so that it can be mapped more cleanly to data. In the model, techniques use a single input, while most real-world producers use multiple. Similarly, α is assumed to be the same for all techniques, but there is enormous variation in cost shares of materials in microdata. Some entrepreneurs may be more capable than others, so that there is a dimension of productivity that is not tied to particular inputs. Finally, as with many studies with heterogeneous producers, it is far from obvious whether the appropriate empirical analog of an entrepreneur is a firm, an establishment, a product, or something else altogether. Further research could help address these gaps.

A key channel in the model is that the network structure—who buys inputs from whom—matters for aggregate productivity. This channel may be useful in assessing the consequences of the marked changes in sourcing during macroeconomic crises or supply chain disruptions.[54] The mechanism may also be important when studying distortions that may cause producers to use the wrong suppliers, leading them to use lower-productivity techniques or higher-cost inputs.[55] Such distortions may include contracting frictions, state mandates to purchase inputs from particular suppliers, or barriers such as the Berlin Wall. Appropriately modified, the model would provide a natural link between such distortions and aggregate productivity.

[54] Gopinath and Neiman (2014) and Lu et al. (2013) documented changes during crises while Carvalho et al. (2016) and Barrot and Sauvagnat (2016) study disruptions following natural disasters.

[55] Jones (2011, 2013) argues that distortions may cause producers to use the wrong input quantities.

References

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza TahbazSalehi, "The Network Origins of Aggregate Fluctuations," *Econometrica*, 09 2012, *80* (5), 1977–2016.

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.

Arkolakis, Costas, "A Unified Theory of Firm Selection and Growth," NBER Working Papers 17553, National Bureau of Economic Research, Inc October 2011.

Atalay, Enghin, "How Important Are Sectoral Shocks?," 2013.

_ , 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.

Axtell, Robert L, "Zipf distribution of US firm sizes," *Science*, 2001, *293* (5536), 1818–1820.

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

Barrot, Jean-Noël and Julien Sauvagnat, "Input specificity and the propagation of idiosyncratic shocks in production networks," *The Quarterly Journal of Economics*, 2016, p. qjw018.

Bingham, NH and RA Doney, "Asymptotic properties of supercritical branching processes I: The Galton-Watson process," *Advances in Applied Probability*, 1974, pp. 711–731.

Buera, Francisco J and Ezra Oberfield, "The Global Diffusion of Ideas," Technical Report, National Bureau of Economic Research 2016.

Caliendo, L. and F. Parro, "Estimates of the Trade and Welfare Effects of NAFTA," Technical Report 2011.

Carvalho, Vasco and Xavier Gabaix, "The great diversification and its undoing," *The American Economic Review*, 2013, *103* (5), 1697–1727.

Carvalho, Vasco M and Nico Voigtländer, "Input diffusion and the evolution of production networks," Technical Report, National Bureau of Economic Research 2014.

_ , Makoto Nirei, Yukiko U Saito, and Alireza Tahbaz-Salehi, "Supply chain disruptions: Evidence from the great east japan earthquake," 2016.

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

Ciccone, Antonio, "Input chains and industrialization," *The Review of Economic Studies*, 2002, *69* (3), 565–587.

de Melo, Rafael Lopes, "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," *Journal of Political Economy*, Forthcoming.

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

Feller, Willliam, *An introduction to probability theory and its applications*, Vol. 2, John Wiley & Sons, 1971.

Foerster, Andrew T., Pierre-Daniel G. Sarte, and Mark W. Watson, "Sectoral versus Aggregate Shocks: A Structural Factor Analysis of Industrial Production," *Journal of Political Economy*, 2011, *119* (1), 1 – 38.

Gabaix, Xavier, "The granular origins of aggregate fluctuations," *Econometrica*, 2011, *79* (3), 733–772.

Gibrat, Robert, *Les inégalités économiques*, Recueil Sirey, 1931.

Giovanni, Julian Di, Andrei A Levchenko, and Isabelle Méjean, "Firms, destinations, and aggregate fluctuations," *Econometrica*, 2014, *82* (4), 1303–1340.

_ , _ , and Romain Ranciere, "Power laws in firm size and openness to trade: Measurement and implications," *Journal of International Economics*, 2011, *85* (1), 42–52.

Gopinath, Gita and Brent Neiman, "Trade adjustment and productivity in large crises," *The American Economic Review*, 2014, *104* (3), 793–831.

Hammond, Peter J, Mamoru Kaneko, and Myrna Holtz Wooders, "Continuum economies with finite coalitions: Core, equilibria, and widespread externalities," *Journal of Economic Theory*, 1989, *49* (1), 113–134.

Hatfield, John William, Scott Duke Kominers, Alexandru Nichifor, Michael Ostrovsky, and Alexander Westkamp, "Stability and competitive equilibrium in trading networks," *Journal of Political Economy*, 2013, *121* (5), 966–1005.

Hopenhayn, Hugo A., "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), pp. 1127–1150.

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

Houthakker, Hendrik, "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis," *The Review of Economic Studies*, 1955, *23* (1), pp. 27–31.

Hulten, Charles R., "Growth Accounting with Intermediate Inputs," *The Review of Economic Studies*, 1978, *45* (3), 511–518.

Jones, Charles I, "Intermediate goods and weak links in the theory of economic development," *American Economic Journal: Macroeconomics*, 2011, *3* (2), 1–28.

_ , "Misallocation, Economic Growth, and Input-Output Economics," in "Advances in Economics and Econometrics: Tenth World Congress," Vol. 2 Cambridge University Press 2013, p. 419.

Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica*, 1982, pp. 649–670.

Jr, Robert E Lucas, "On the size distribution of business firms," *The Bell Journal of Economics*, 1978, pp. 508–523.

Kelly, Bryan, Hanno Lustig, and Stijn Van Nieuwerburgh, "Firm Volatility in Granular Networks," 2013.

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

Klette, T.J. and S. Kortum, "Innovating firms and aggregate innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

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

Kramarz, Francis, Julien Martin, and Isabelle Mejean, "Volatility in the Small and in the Large: The Lack of Diversification in International Trade," 2016.

Leontief, Wassily W, "Input-Output Economics," *Scientific American*, 1951, *185* (4), 15–21.

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

Lu, Dan, Asier Mariscal, and Luis Fernando Mejia, "Imports Switching and the Impact of Large Devaluation," Technical Report 2013.

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

Lucas, Robert E and Benjamin Moll, "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 2014, *122* (1).

Luttmer, Erzo G. J., "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, 08 2007, *122* (3), 1103–1144.

Pavitt, Keith, "Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory," *Research policy*, 1984, *13* (6), 343–373.

Perla, Jesse and Christopher Tonetti, "Equilibrium imitation and growth," *Journal of Political Economy*, 2014, *122* (1), 52–76.

Rosen, Sherwin, "The economics of superstars," *The American economic review*, 1981, pp. 845–858.

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

Rossi-Hansberg, Esteban and Mark L. J. Wright, "Establishment Size Dynamics in the Aggregate Economy," *American Economic Review*, December 2007, *97* (5), 1639–1666.

Scherer, Frederick M, "Inter-Industry Technology Flows in the United States," *Research Policy*, 1982, *11* (4), 227–245.

Simon, Herbert A, "On a class of skew distribution functions," *Biometrika*, 1955, *42* (3/4), 425–440.

Simon, Herbert and Charles P. Bonini, "The Size Distribution of Business Firms," *The American Economic Review*, 1958, *48* (4), 607–617.

Timmer, Marcel P, Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J Vries, "An illustrated user guide to the world input–output database: the case of global automotive production," *Review of International Economics*, 2015, *23* (3), 575–605.

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

Weitzman, M.L., "Recombinant growth," *The Quarterly Journal of Economics*, 1998, *113* (2), 331–360.

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

Yule, G Udny, "A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FRS," *Philosophical transactions of the Royal Society of London. Series B, containing papers of a biological character*, 1925, *213*, 21–87.

Online Appendix

A Stable Equilibria

A.1 Notation

This subsection shows that it is without loss of generality to associate each bilateral contract with a technique and restrict the arrangement so that there is one contract for each technique.

A coalition is *connected* if for any entrepreneurs i and j in the coalition, there is a sequence of entrepreneurs beginning with i and ending with j, all of whom are members of the coalition, such that for each consecutive pair there exists a technique in Φ for which one entrepreneur is the buyer and the other is the supplier.

Lemma 1. *If there is a coalition of size/cardinality n or smaller with a dominating deviation, then there is a connected coalition of size/cardinality n or smaller with a dominating deviation.*

Proof. Suppose that there is a coalition J with a dominating deviation that can be divided into two subsets that are not connected, J' and J'', so that $J' \cup J'' = J$ and $J' \cap J'' = \varnothing$. The original deviation would leave every member of J at least as well off and at least one member of J strictly better off. Without loss of generality, suppose that the member who is strictly better off is in J'. Then J' has a dominating deviation: Since no member of J' is able to produce using intermediate inputs from members of J'', J' has a dominating deviation in which members of J' drop all contracts for which there are positive payments to members of J''; every member of J' is at least as well off as under the original deviation. ■

Lemma 2. *It is without loss of generality to use notation that associates each bilateral contract with a technique.*

Proof. We first show if j has no technique to use i's good as an input, then there is no equilibrium in which i provides goods to j or there is a payment between them. After that, we will show that if a coalition has a dominating deviation, then there is always an alternative dominating deviation in which each pairwise payment and trade of goods can be associated with a technique.

Toward an contradiction, suppose first there is an equilibrium in which entrepreneur i provides goods for entrepreneur j, but j has no technique that would use good i as an input. Since j cannot resell good i, if the payment to j is positive then i would be strictly better off dropping the contract (setting the payment and the quantity of goods to zero). If the payment is weakly negative, i would be strictly better off dropping the contract. Thus this cannot be an equilibrium. Suppose second that there is an equilibrium in which there is a payment from i to j but no goods are provided. Then unless the payment is zero, one of the two would be strictly better off dropping the contract.

Next, suppose the connected coalition J has a dominating deviation in which there is a payment from i to j. Then because the coalition is connected, there is another deviation with identical payoffs where the payment from i to j is payoff is intermediated by those on the path from i to j. ■

Lemma 3. *It is without loss of generality to use notation that associates each technique with a single bilateral contract.*

Proof. Suppose there is an arrangement in which there may be multiple bilateral contracts are associated with each technique. Suppose further that the arrangement is stable with respect to deviations by coalitions of size/cardinalty n. Then the alternative arrangement in which all contracts for each technique are combined into a single contract delivers the same allocation and payoffs and must also be stable with respect to deviations by coalitions of size/cardinalty n because those deviations were available for the original arrangement. ■

A.2 A Supply Chain Representation

This section describes notation that decomposes the allocation into production within the many supply chains available to produce the various goods. Such a mapping can be constructed because each technique exhibits constant returns to scale.

For a supply chain $\mathfrak{c} \in \mathfrak{C}_j$ available to produce good j, we can summarize production at each step in the chain in the eventual production of j for final consumption. Towards this, we will construct the variables $\{y^n(\mathfrak{c}), x^n(\mathfrak{c}), l^n(\mathfrak{c})\}$ to be the quantities of output, intermediate input, and labor used in the n^{th}-to-last step in the supply chain $\mathfrak{c}$ in the production of good j for final consumption.

Entrepreneur j's output using technique $\phi \in U_j$ is $y(\phi)$. j's output used as an intermediate input for technique $\phi' \in D_j$ is $x(\phi')$, and its output for the household's consumption is y_j^0. For $\phi' \in D_j$ and $\phi \in U_j$, let $\Upsilon_{\phi'}(\phi)$ be the quantity of $x(\phi')$ that is made using technique ϕ, and let $\Upsilon_0(\phi)$ be the quantity of y_j^0 that is made using technique ϕ. Therefore $\sum_{\phi \in U_j} \Upsilon_{\phi'}(\phi) = x(\phi')$, $\sum_{\phi \in U_j} \Upsilon_0(\phi) = y_j^0$, and $\Upsilon_0(\phi) + \sum_{\phi' \in D_j} \Upsilon_{\phi'}(\phi) = y(\phi)$.

Consider a supply chain $\mathfrak{c} \in \mathfrak{C}_j$. Let $\phi^n(\mathfrak{c})$ denote the n^{th} technique in the chain so that, for example, $\phi^0(\mathfrak{c})$ is the most downstream technique, and let $\mathfrak{j}^n(\mathfrak{c})$ be the identity of the buyer associated with that technique so that $\mathfrak{j}^0(\mathfrak{c}) = j$ and $\mathfrak{j}^n(\mathfrak{c}) = b(\phi^n(\mathfrak{c})) = s(\phi^{n-1}(\mathfrak{c}))$. In the representation, let $y^0(\mathfrak{c})$ be the total amount of good j produced for consumption using supply chain $\mathfrak{c}$.[56] Thus total consumption of good j is $y_j^0 = \sum_{\mathfrak{c}\in\mathfrak{C}_j} y^0(\mathfrak{c})$. $y^0(\mathfrak{c})$ is

$$y^0(\mathfrak{c}) \equiv \Upsilon_0\left(\phi^0(\mathfrak{c})\right) \prod_{k=0}^{\infty} \frac{\Upsilon_{\phi^k(\mathfrak{c})}\left(\phi^{k+1}(\mathfrak{c})\right)}{x_{\phi^k(\mathfrak{c})}}$$

$\Upsilon_0\left(\phi^0(\mathfrak{c})\right)$ the output of good j from technique $\phi^0(\mathfrak{c})$ that goes to the household for consumption, while $\frac{\Upsilon_{\phi^k(\mathfrak{c})}\left(\phi^{k+1}(\mathfrak{c})\right)}{x_{\phi^k(\mathfrak{c})}}$ is the fraction of production using technique $\phi^k(\mathfrak{c})$ that is produced using technique $\phi^{k+1}(\mathfrak{c})$.[57]

We next construct $\left\{l^n(\mathfrak{c}), x^n(\mathfrak{c}), y^{n+1}(\mathfrak{c})\right\}_{n=0}^{\infty}$ iteratively using the following equalities:

$$\frac{y^n(\mathfrak{c})}{y\left(\phi^n(\mathfrak{c})\right)} = \frac{l^n(\mathfrak{c})}{l\left(\phi^n(\mathfrak{c})\right)} = \frac{x^n(\mathfrak{c})}{x\left(\phi^n(\mathfrak{c})\right)}$$

along with $x^n(\mathfrak{c}) = y^{n+1}(\mathfrak{c})$. The first two equalities simply indicate that the fraction of output of the n^{th} technique that is used for the n^{th}-to-last stage of production of good j for consumption equals the corresponding fractions of labor and intermediate inputs. The final equality simply states that the output at one stage in a chain is the intermediate input used in the subsequent stage.

It will be useful to define the efficiency of a supply chain to be $q(\mathfrak{c}) \equiv \prod_{n=0}^{\infty} z^n(\mathfrak{c})^{\alpha^n}$. In general, $\lim_{N\to\infty} \prod_{n=0}^{N-1} z^n(\mathfrak{c})^{\alpha^n}$ may not exist. However, the following claim shows that under Assumption 1 the limit always exists.

Lemma 4. *Assume $z_0 > 0$. Then for each $\mathfrak{c} \in \mathfrak{C}_j$, $\lim_{N\to\infty} \prod_{n=0}^{N-1} z^n(\mathfrak{c})^{\alpha^n}$ exists.*

Proof. For each n, $z^n(\mathfrak{c})$ can be decomposed into $z^{n+}(\mathfrak{c})z^{n-}(\mathfrak{c})$, where $z^{n+}(\mathfrak{c}) = \max\{z^n(\mathfrak{c}), 1\}$ and $z^{n-}(\mathfrak{c}) = \min\{z^n(\mathfrak{c}), 1\}$, so that $\prod_{n=0}^{N-1} z^n(\mathfrak{c})^{\alpha^n} = \left(\prod_{n=0}^{N-1} z^{n+}(\mathfrak{c})^{\alpha^n}\right)\left(\prod_{n=0}^{N-1} z^{n-}(\mathfrak{c})^{\alpha^n}\right)$. $\prod_{n=0}^{N-1} z^{n+}(\mathfrak{c})^{\alpha^n}$ is a monotone sequence so it converges to a (possibly infinite) limit. $\prod_{n=0}^{N-1} z^{n-}(\mathfrak{c})^{\alpha^n}$ is a monotone sequence bounded below by $z_0^{\frac{1}{1-\alpha}}$ so it converges to a limit in the range $\left[z_0^{\frac{1}{1-\alpha}}, 1\right]$. Thus

[56] This is distinct from the total quantity of good j produced using the supply chain $\mathfrak{c}$; the latter includes production for intermediate use by other entrepreneurs.

[57] It will be shown later that in equilibrium, generically each entrepreneur uses only single technique. This means that generically an entrepreneur's output will be produced using a single supply chain; for the supply chain that is actually used $y^0(\mathfrak{c}) = y_j^0$, while for the supply chains that are not used $y^0(\mathfrak{c}) = 0$.

$\prod_{n=0}^{N-1} z^n(\mathfrak{c})^{\alpha^n}$ converges to a (possibly infinite) limit. ■

A.3 Proof of Proposition 1

Lemma 5. *For any arrangement, the following hold*

$$
\begin{aligned}
p_j^0 &= \frac{\varepsilon}{\varepsilon-1}\lambda_j \\
\frac{l(\phi)}{1-\alpha} &= \left[\frac{\lambda_j z(\phi)}{w}\right]^{\frac{1}{\alpha}} \frac{x(\phi)}{\alpha} && (16)\\
\lambda_j &= w\left[\frac{y_j^0+\sum_{\phi\in D_j} x(\phi)}{\sum_{\phi\in U_j}\frac{1}{\alpha}z(\phi)^{\frac{1}{\alpha}}x(\phi)}\right]^{\frac{\alpha}{1-\alpha}} && (17)\\
l_j &= (1-\alpha)\left[y_j^0+\sum_{\phi\in D_j} x(\phi)\right]^{\frac{1}{1-\alpha}}\left[\sum_{\phi\in U_j}\frac{1}{\alpha}z(\phi)^{\frac{1}{\alpha}}x(\phi)\right]^{-\frac{\alpha}{1-\alpha}} && (18)
\end{aligned}
$$

Proof. Together, the FOCs with respect to p_j^0 and y_j^0 imply $p_j^0 = \frac{\varepsilon}{\varepsilon-1}\lambda_j$. Individual rationality guarantees that $x(\phi)=0$ implies $l(\phi)=0$. If $x(\phi)>0$, the FOC with respect to $l(\phi)$ is $w = \lambda_j \frac{1-\alpha}{\alpha^\alpha(1-\alpha)^{1-\alpha}} z(\phi)x(\phi)^\alpha l(\phi)^{-\alpha}$. which can be rearranged as

$$
l(\phi) = \left[\frac{\lambda_j z(\phi)}{w}\right]^{\frac{1}{\alpha}} \frac{1-\alpha}{\alpha} x(\phi)
$$

Substituting this into the entrepreneur's constraint and solving for λ_j

$$
\begin{aligned}
y_j^0 + \sum_{\phi\in D_j} x(\phi) &= \sum_{\phi\in U_j}\frac{1}{\alpha^\alpha(1-\alpha)^{1-\alpha}} z(\phi)x(\phi)^\alpha\left[\frac{\lambda_j\frac{1-\alpha}{\alpha^\alpha(1-\alpha)^{1-\alpha}} z\left(\phi\right)x(\phi)^\alpha}{w}\right]^{\frac{1-\alpha}{\alpha}}\\
\frac{\lambda_j}{w} &= \left[\frac{y_j^0+\sum_{\phi\in D_j} x(\phi)}{\sum_{\phi\in U_j}\frac{1}{\alpha}z(\phi)^{\frac{1}{\alpha}}x(\phi)}\right]^{\frac{\alpha}{1-\alpha}}
\end{aligned}
$$

The first order condition for labor can thus be written as

$$
1 = \left[\frac{y_j^0+\sum_{\phi\in D_j} x(\phi)}{\sum_{\phi\in U_j}\frac{1}{\alpha}z(\phi)^{\frac{1}{\alpha}}x(\phi)}\right]^{\frac{\alpha}{1-\alpha}} \frac{1-\alpha}{\alpha^\alpha(1-\alpha)^{1-\alpha}} z(\phi)x(\phi)^\alpha l(\phi)^{-\alpha}
$$

Solving for $l(\phi)$

$$l(\phi) = \frac{(1-\alpha)}{\alpha}\left[\frac{y_j^0+\sum_{\phi\in D_j}x(\phi)}{\sum_{\phi\in U_j}\frac{1}{\alpha}z(\phi)^{\frac{1}{\alpha}}x(\phi)}\right]^{\frac{1}{1-\alpha}} z(\phi)^{\frac{1}{\alpha}}x(\phi)$$

Summing across techniques yields (18). ■

Lemma 6. *If an arrangement is pairwise stable, then for each $\phi\in\Phi$, $z(\phi)q_{s(\phi)}^{\alpha}\leq q_{b(\phi)}$ with equality if $x(\phi)>0$*

Proof. Note first that (17) implies that if $q_j=0$, then it must be that $y_j=0$. We now proceed in three cases. First, if $q_{s(\phi)}=0$ then the statement is true because $x(\phi)=0$ and $0\leq q_{b(\phi)}$.

Second, suppose that $q_{s(\phi)}>0$ and $q_{b(\phi)}>0$. The envelope theorem implies that, to a first order, an increase in $x(\phi)$ reduces $s(\phi)$'s profit by $\lambda_{s(\phi)}$. To assess the impact on the buyer's profit, it will be useful to plug in (18) to the buyer's problem so it can be written as

$$\pi_j = \max_{p_j^0,y_j^0} p_j^0y_j^0 - w(1-\alpha)\left[y_j^0+\sum_{\phi\in D_j}x(\phi)\right]^{\frac{1}{1-\alpha}}\left[\sum_{\phi\in U_j}\frac{1}{\alpha}z(\phi)^{\frac{1}{\alpha}}x(\phi)\right]^{-\frac{\alpha}{1-\alpha}} - \sum_{\phi\in U_j}T(\phi)+\sum_{\phi\in D_j}T(\phi)$$

subject to $y_j^0\leq\left(p_j^0/P^0\right)^{-\varepsilon}Y$. Then the envelope theorem implies that, to a first order, an increase in $x(\phi)$ raises $b(\phi)$'s profit by

$$w\left[y_j^0+\sum_{\phi\in D_j}x(\phi)\right]^{\frac{1}{1-\alpha}}\left[\sum_{\phi\in U_j}\frac{1}{\alpha}z(\phi)^{\frac{1}{\alpha}}x(\phi)\right]^{-\frac{\alpha}{1-\alpha}-1} z(\phi)^{\frac{1}{\alpha}}$$

Using (17) this equals

$$w\left(\frac{\lambda_{b(\phi)}}{w}\right)^{\frac{1}{\alpha}}z(\phi)^{\frac{1}{\alpha}}$$

Pairwise stability implies that there can be no better contract between $b(\phi)$ and $s(\phi)$. If $x(\phi)>0$, there can be no gains from either increasing or reducing $x(\phi)$, which implies $w\left(\frac{\lambda_{b(\phi)}}{w}\right)^{\frac{1}{\alpha}}z(\phi)^{\frac{1}{\alpha}}=\lambda_{s(\phi)}$, or $q_{b(\phi)}=z(\phi)q_{s(\phi)}^{\alpha}$. If $x(\phi)=0$, there can be no gains from increasing $x(\phi)$, so it must be that $w\left(\frac{\lambda_{b(\phi)}}{w}\right)^{\frac{1}{\alpha}}z(\phi)^{\frac{1}{\alpha}}\leq\lambda_{s(\phi)}$, or $q_{b(\phi)}\geq z(\phi)q_{s(\phi)}^{\alpha}$.

Finally, suppose that $q_{s(\phi)}>0$ but $q_{b(\phi)}=0$. The latter implies that $y_{b(\phi)}=0$. Consider a deviation in which $\tilde{x}(\phi)=\eta$. If the buyer chooses (suboptimally) to use $\tilde{l}(\phi)=\frac{1-\alpha}{\alpha}\eta\frac{\lambda_{s(\phi)}}{w}$ units of labor with the technique, its output would be $\frac{1}{\alpha}\frac{z(\phi)}{q_{s(\phi)}^{1-\alpha}}\eta$. Since $y_{b(\phi)}=0$, the deviation implies $\tilde{y}^0_{b(\phi)} = \frac{1}{\alpha}\frac{z(\phi)}{q^{1-\alpha}_{s(\phi)}}\eta$ and $\tilde{p}^0_{b(\phi)} = P\left(\tilde{y}^0_{b(\phi)}/Y\right)^{-\frac{1}{\varepsilon}}$. $\tilde{p}^0_{b(\phi)}\tilde{y}^0_{b(\phi)}$ is proportional to $\eta^{\frac{\varepsilon-1}{\varepsilon}}$ while $w\tilde{l}(\phi)$ is proportional to η. For η small enough, the cost to the supplier is of order η. Thus there is a contract with η small enough that would increase the joint surplus. ■

Lemma 7. *Pairwise stability implies that for all ϕ, $\frac{wl(\phi)}{1-\alpha} = \frac{\lambda_{s(\phi)}x(\phi)}{\alpha} = \lambda_{b(\phi)}y(\phi)$ and for each j,*

$$(1-\alpha)\lambda_j y_j = \sum_{\phi\in U_j} wl(\phi) \tag{19}$$

$$\alpha\lambda_j y_j = \sum_{\phi\in U_j} \lambda_{s(\phi)}x(\phi) \tag{20}$$

Proof. Individual rationality guarantees that $x(\phi) = 0$ implies $l(\phi) = 0$. If $x(\phi) > 0$, then individual rationality implies $\frac{l(\phi)}{1-\alpha} = \left[\frac{\lambda_{b(\phi)}z(\phi)}{w}\right]^{\frac{1}{\alpha}}\frac{x(\phi)}{\alpha}$ and pairwise stability implies $z(\phi)q^{\alpha}_{s(\phi)} = q_{b(\phi)}$. Together, these imply $\frac{wl(\phi)}{1-\alpha} = \frac{\lambda_{s(\phi)}x(\phi)}{\alpha}$. These along with the production function imply $\frac{\lambda_{s(\phi)}x(\phi)}{\alpha} = \lambda_{b(\phi)}y(\phi)$. Summing over techniques $\phi\in U_j$ gives (19) and (20). ■

With this in hand we turn to the implications of stability for entrepreneurs' efficiencies. Proposition 1 gives the implications of pairwise and countable stability.

Proposition 1 (1 of 6). *In any pairwise stable equilibrium, if U_j is non-empty then $q_j = \max_{\phi\in U_j} z(\phi)q^{\alpha}_{s(\phi)}$.*

Proof. This follows immediately from Lemma 6. ■

Proposition 1 (2 of 6). *In any pairwise stable equilibrium, $Y^0 = QL$ and $y^0_j = q^{\varepsilon}_j(Q)^{1-\varepsilon}L$.*

Proof. For each chain $\mathfrak{c}\in\mathfrak{C}_j$, it must be that $\lambda_j y^0(\mathfrak{c}) = (1-\alpha)wl^0(\mathfrak{c})$, $(1-\alpha)wl^n(\mathfrak{c}) = \alpha\lambda_{\mathfrak{j}^{n+1}(\mathfrak{c})}x^n(\mathfrak{c}) = \alpha(1-\alpha)wl^{n+1}(\mathfrak{c})$ (these are true whether or not $y^0(\mathfrak{c})$, $l^n(\mathfrak{c})$, and $x^n(\mathfrak{c})$ are positive). Together, these imply that $l^n(\mathfrak{c}) = \alpha^n(1-\alpha)\frac{1}{q_j}y^0(\mathfrak{c})$. Using $p^0_j = \frac{\varepsilon}{\varepsilon-1}\frac{w}{q_j}$ and hence $P^0 = \left(\int_0^1 (p^0_j)^{1-\varepsilon}dj\right)^{\frac{1}{1-\varepsilon}} = \frac{\varepsilon}{\varepsilon-1}\frac{w}{Q}$, the quantity sold to the household is $y^0_j = (q_j/Q)^{\varepsilon}Y^0$.

Summing over the labor used in all steps in each supply chain in $\mathfrak{C}_j$ we have

$$\begin{aligned} L &= \int_0^1 \sum_{\mathfrak{c}\in\mathfrak{C}_j}\sum_{n=0}^{\infty} l^n(\mathfrak{c})dj = \int_0^1 \sum_{\mathfrak{c}\in\mathfrak{C}_j}\sum_{n=0}^{\infty}\alpha^n(1-\alpha)\frac{1}{q_j}y^0(\mathfrak{c})dj \\ &= \int_0^1 \frac{1}{q_j}y^0_j dj = \int_0^1 \frac{1}{q_j}(q_j/Q)^{\varepsilon}Y^0 dj = \frac{1}{Q}Y^0 \end{aligned}$$

Plugging this back into $y^0_j = (q_j/Q)^{\varepsilon}Y^0$ gives the result. ■

Proposition 1 (3 of 6). *In any pairwise stable equilibrium, $q_j \leq \sup_{\mathfrak{c}\in\mathfrak{C}_j} q(\mathfrak{c})$.*

Proof. If $q_j = 0$, the conclusion is immediate. Suppose not. Pairwise stability implies that

$$\begin{aligned}
\lambda_{\mathfrak{j}^{n+1}(\mathfrak{c})}x^n(\mathfrak{c}) &= \alpha\lambda_{\mathfrak{j}^n(\mathfrak{c})}y^n(\mathfrak{c}) \\
wl^n(\mathfrak{c}) &= (1-\alpha)\lambda_{\mathfrak{j}^n(\mathfrak{c})}y^n(\mathfrak{c}) \\
y^n(\mathfrak{c}) &= x^{n-1}(\mathfrak{c}) \\
z^n(\mathfrak{c})\left(\frac{w}{\lambda_{\mathfrak{j}^{n+1}(\mathfrak{c})}}\right)^\alpha &\leq \frac{w}{\lambda_{\mathfrak{j}^n(\mathfrak{c})}}
\end{aligned}$$

Together these imply that for any chain $\mathfrak{c} \in \mathfrak{C}_j$,

$$\begin{aligned}
\frac{wl^n(\mathfrak{c})}{1-\alpha} &= \lambda_{\mathfrak{j}^n(\mathfrak{c})}y^n(\mathfrak{c}) = \lambda_{\mathfrak{j}^n(\mathfrak{c})(\mathfrak{c})}x^{n-1}(\mathfrak{c}) = \alpha\lambda_{\mathfrak{j}^{n-1}(\mathfrak{c})}y^{n-1}(\mathfrak{c}) = \alpha\lambda_{\mathfrak{j}^{n-1}(\mathfrak{c})}x^{n-2}(\mathfrak{c}) = \alpha^2\lambda_{\mathfrak{j}^{n-2}(\mathfrak{c})}y^{n-2}(\mathfrak{c}) \\
&= \alpha^n\lambda_{\mathfrak{j}^0(\mathfrak{c})}y^0(\mathfrak{c})
\end{aligned}$$

or $\frac{l^n(\mathfrak{c})}{1-\alpha} = \alpha^n\frac{1}{q_j}y^0(\mathfrak{c})$. Plugging these into the chain feasibility condition for chain $\mathfrak{c} \in \mathfrak{C}_j$ gives

$$\begin{aligned}
y^0(\mathfrak{c}) &\leq \prod_{n=0}^{\infty}\left(\frac{1}{\alpha^\alpha(1-\alpha)^{1-\alpha}}z^n(\mathfrak{c})l^n(\mathfrak{c})^{1-\alpha}\right)^{\alpha^n} \\
&\leq \prod_{n=0}^{\infty}\left(\frac{1}{\alpha^\alpha}z^n(\mathfrak{c})\left\{\alpha^n\frac{1}{q_j}y^0(\mathfrak{c})\right\}^{1-\alpha}\right)^{\alpha^n}
\end{aligned}$$

Note that $\prod_{n=0}^{\infty}\left(\frac{\{\alpha^n\}^{1-\alpha}}{\alpha^\alpha}\right)^{\alpha^n} = \frac{\prod_{n=0}^{\infty}\left(\alpha^{1-\alpha}\right)^{n\alpha^n}}{\prod_{n=0}^{\infty}(\alpha^\alpha)^{\alpha^n}} = \frac{\alpha^{1-\alpha\sum_{n=0}^{\infty}n\alpha^n}}{\alpha^{\alpha\sum_{n=0}^{\infty}\alpha^n}} = \frac{\alpha^{(1-\alpha)\frac{\alpha}{(1-\alpha)^2}}}{\alpha^{\alpha\frac{1}{1-\alpha}}} = 1$. Using $q(\mathfrak{c}) \equiv \prod_{n=0}^{\infty} z^n(\mathfrak{c})^{\alpha^n}$, the chain feasibility condition becomes

$$y^0(\mathfrak{c}) \leq q(\mathfrak{c})\prod_{n=0}^{\infty}\left(\left\{\frac{1}{q_j}y^0(\mathfrak{c})\right\}^{1-\alpha}\right)^{\alpha^n} = \frac{q(\mathfrak{c})}{q_j}y^0(\mathfrak{c}) \tag{21}$$

Towards a contradiction, suppose that $\eta q_j = \sup_{\mathfrak{c}\in\mathfrak{C}_j} q(\mathfrak{c})$ with $\eta < 1$. Then (21) implies $y^0(\mathfrak{c}) \leq \eta y^0(\mathfrak{c})$. Summing across all chains $\mathfrak{c} \in \mathfrak{C}_j$ gives $y_j^0 \leq \eta y_j^0$. $q_j > 0$ implies $y_j^0 > 0$, and hence we have a contradiction. ■

Proposition 1 (4 of 6). *In any countably stable arrangement, for each j, $q_j = \sup_{\mathfrak{c}\in\mathfrak{C}_j} q(\mathfrak{c})$.*

Proof. Toward a contradiction, suppose there is a j and a $\mathfrak{c} \in \mathfrak{C}_j$ such that $q(\mathfrak{c}) > q_j$. For any integer $n \geq 0$, define $q_n(\mathfrak{c}) \equiv \prod_{k=n}^{\infty} z^k(\mathfrak{c})^{\alpha^k}$ so that $q_0(\mathfrak{c}) = q(\mathfrak{c})$ and $q_1(\mathfrak{c})$ is the maximum feasible efficiency for the next to last entrepreneur in the chain $\mathfrak{c}$ that is consistent with pairwise stability. Since the arrangement is pairwise stable, j's total spending on labor is $wl_j = \sum_{\phi \in U_j} wl(\phi) = (1-\alpha)\lambda_j y_j$.

We will show that there is a countable coalition with a dominating deviation. The deviation has two parts. Let $\eta \in (0,1)$. For the first part, j lowers its spending on labor used with each technique, so that $\tilde{l}(\phi) = \eta^{\frac{1}{1-\alpha}} l(\phi)$. This reduces j's spending on wages by $\left(1 - \eta^{\frac{1}{1-\alpha}}\right) \lambda_j y_j$ and reduces its output to ηy_j.

For the second part, the entire supply chain $\mathfrak{c}$ increases production to make up for j's lost output. The deviation will leave each of the suppliers in the supply chain equally well off, but j better off. For each integer $n \geq 0$, let $\tilde{x}^n(\mathfrak{c})$ and $\tilde{T}^n(\mathfrak{c})$ satisfy $\left[\tilde{x}^n(\mathfrak{c}) - x^n(\mathfrak{c})\right] \lambda_{\mathrm{j}^{n+1}(\mathfrak{c})} = \alpha^{n+1} \lambda_{\mathrm{j}^0(\mathfrak{c})}(1-\eta) y_j$ and $\tilde{T}^n(\mathfrak{c}) = T^n(\mathfrak{c}) + \left[\tilde{x}^n(\mathfrak{c}) - x^n(\mathfrak{c})\right] \lambda_{\mathrm{j}^{n+1}(\mathfrak{c})}$. For $n \geq 1$, The entrepreneur at the n^{th} step in the chain could attain the same payoff as before the deviation choosing $\tilde{l}^n(\mathfrak{c}) = l^n(\mathfrak{c}) + \alpha^n(1-\alpha)(1-\eta) y_j \frac{\lambda_{\mathrm{j}^0(\mathfrak{c})}}{w}$, $\tilde{p}^0_{\mathrm{j}^n(\mathfrak{c})} = p^0_{\mathrm{j}^n(\mathfrak{c})}$, $\tilde{y}^0_{\mathrm{j}^n(\mathfrak{c})} = y^0_{j^n(\mathfrak{c})}$. Entrepreneur j on the other hand could increase labor by $(1-\alpha)\lambda_{\mathrm{j}^0(\mathfrak{c})}(1-\eta) y_j$.

For j, the cost savings from the first part would be $\left(1 - \eta^{\frac{1}{1-\alpha}}\right) \lambda_j y_j$, while the increased spending from the second part would be $(1-\eta) y_j \lambda(\mathfrak{c})$ (which accounts for both the increased spending on labor and the payment to the supplier). The change in j's payoff is thus

$$\left(1 - \eta^{\frac{1}{1-\alpha}}\right) \lambda_j y_j - (1-\eta) y_j \lambda(\mathfrak{c})$$

For η close enough to 1, this is strictly positive. ■

The next two claims study efficiency and uniqueness of allocations consistent with countable stability.

Proposition 1 (5 of 6)**.** *Every countably stable equilibrium is efficient.*

Proof. Consider the problem of a planner that, taking the set of techniques Φ as given, makes production decisions and allocates labor to maximize the utility of the representative consumer. For each producer $j \in [0,1]$, the planner chooses the quantity to be produced for final consumption, y^0_j. In addition, for each of j's upstream techniques $\phi \in U_j$, the planner chooses a quantity of labor, $l(\phi)$, and a quantity of good $s(\phi)$ for j to use as an intermediate input, $x(\phi)$.

Alternatively, we can formulate the planner's problem with the supply chain representation: For each good j and for each supply chain $\mathfrak{c} \in \mathfrak{C}_j$, the planner chooses the labor, intermediate inputs and output used at each step in those chains to produce consumption of good j. Following the logic of Section 2.2, if the planner used supply chain $\mathfrak{c} \in \mathfrak{C}_j$ to produce good j for consumption, its indirect production function would be $y^0(\mathfrak{c}) = q(\mathfrak{c})l$. Since the planner would choose to produce good j in the least cost way possible, it would use the most efficient supply chain, so that $y_j^0 = q_j^{\text{planner}} \bar{l}_j$, where $q_j^{\text{planner}} = \sup_{\mathfrak{c} \in \mathfrak{C}_j} q(\mathfrak{c})$ and $\bar{l}_j$ is the total labor the planner uses across all steps in all chains to produce good j for the household. Thus the planner's problem can be restated as $\max_{Y^0, \{y_j^0, \bar{l}_j\}_{j\in[0,1]}} Y^0$ subject to $Y^0 = \left(\int_0^1 \left(y_j^0\right)^{\frac{\varepsilon-1}{\varepsilon}} dj\right)^{\frac{\varepsilon}{\varepsilon-1}}$ and $y_j^0 \leq q_j^{\text{planner}} \bar{l}_j, \forall j$. This yields $Y^0 = Q^{\text{planner}} L$, where $Q^{\text{planner}} = \left(\int_0^1 (q_j^{\text{planner}})^{\varepsilon-1} dj\right)^{\frac{1}{\varepsilon-1}}$. In any countably-stable equilibrium, $q_j = q_j^{\text{planner}}$, so it must be that $Q = Q^{\text{planner}}$ and the equilibrium is efficient. ■

Given Φ, there may be multiple allocations consistent with countable stability. Let $J^{unique}(\Phi)$ be the set of entrepreneurs for whom all production variables (i.e., $\{x(\phi), l(\phi), y(\phi)\}_{\phi \in U_j}$, $\{x(\phi)\}_{\phi \in D_j}$, p_j^0, y_j^0) are the same across all countably-stable equilibria. The

Proposition 1 (6 of 6)**.** *Suppose H is atomless. Then with probability one, Φ is such that $J^{unique}(\Phi)$ has unit measure.*

Proof. We first show that the probability that an entrepreneur has two techniques that deliver the same efficiency is zero. This is follows from the fact that, given each potential supplier's efficiency, the efficiency delivered by the technique is $z(\phi) q_{s(\phi)}^{\alpha}$. Since H is atomless, the probability that any finite set of techniques has two that deliver the same efficiency is zero. Second, the probability of entrepreneur j or any entrepreneur downstream from j has two techniques that deliver the same efficiency is zero. to see this, note that since the number of downstream techniques is countable, they can be ordered. For any N, the probability that any of the first N entrepreneurs have two techniques that deliver the same efficiency is 0^N. Thus the probability that any downstream entrepreneur has two such techniques is $\lim_{N\to\infty} 0^N = 0$. Third, if no entrepreneurs downstream from j have two techniques that deliver the same efficiency, then $j \in J^{unique}(\Phi)$. To see this, consider some entrepreneur $\tilde{j}$ that is downstream from j. $\tilde{j}$'s sells $y_{\tilde{j}}^0 = q_{\tilde{j}}^{\varepsilon-1} Q^{\varepsilon} Y^0$ to the household for consumption. If j is the n^{th} supplier in $\tilde{j}$'s best supply chain, then in every countably-stable equilibrium j produces $\frac{1}{\lambda_j}\alpha^n \lambda_{\tilde{j}} y_{\tilde{j}}^0$ to be used in the supply chain to produce $\tilde{j}$ for consumption, using $\frac{1}{w}(1-\alpha)\alpha^n \lambda_{\tilde{j}} y_{\tilde{j}}^0$ units of labor and $\frac{1}{\lambda_s}\alpha^{n+1}\lambda_{\tilde{j}} y_{\tilde{j}}^0$ units of intermediate inputs (where λ_s is the marginal cost of the supplier used by j). And, of course, if j is not in $\tilde{j}$'s best supply chain, all of these quantities are zero. Total labor, intermediate inputs and output of entrepreneur j is simply the sum of these quantities over all entrepreneurs weakly downstream from j.

■

Lemma 8. *If an arrangement generates an allocation such that for each $j \in [0,1]$, $q_j = \max_{\phi \in U_j} z(\phi) q_{s(\phi)}^{\alpha}$ and $q_j \leq sup_{\mathfrak{c} \in \mathfrak{C}_j} q(\mathfrak{c})$, the allocation is feasible.*

Proof. Given that the market for each good clears (this is built into the constraints of the entrepreneurs' problems) and the households budget constraint is satisfied, Walras' law implies that the wage is such that the labor market clears as well. Thus resource feasibility is satisfied and we only need to verify chain feasibility.

j's marginal cost is $\lambda_j = w/q_j$, so the price it charges the household is $\frac{\varepsilon}{\varepsilon-1}\frac{w}{q_j}$. The total cost of labor used by j across all chains for j to produce output for the household is $\sum_{\mathfrak{c} \in \mathfrak{C}_j} wl^0(\mathfrak{c}) = (1-\alpha)\sum_{\mathfrak{c} \in \mathfrak{C}_j} \lambda_j y^0(\mathfrak{c})$.

We next show that for any chain $\mathfrak{c} \in \mathfrak{C}_j$, $l^n(\mathfrak{c}) = \alpha l^{n-1}(\mathfrak{c})$. To see this, note that for any technique, individual rationality implies that

$$\frac{wl(\phi)}{1-\alpha} = w^{1-1/\alpha}\left[\lambda_{b(\phi)} z(\phi)\right]^{1/\alpha} \frac{x(\phi)}{\alpha}$$

We also have that $\lambda_{b(\phi)} \leq \frac{1}{z(\phi)} w^{1-\alpha} \lambda_{s(\phi)}^{\alpha}$ with equality if $x(\phi) > 0$. Whether or not $x(\phi) > 0$, these imply

$$\frac{wl(\phi)}{1-\alpha} = \frac{\lambda_{s(\phi)} x(\phi)}{\alpha}$$

This implies that in the supply chain representation, for any $\mathfrak{c}$ and n,

$$\frac{wl^n(\mathfrak{c})}{1-\alpha} = \frac{\lambda_{\mathfrak{j}^{n+1}(\mathfrak{c})} x^n(\mathfrak{c})}{\alpha}$$

Since $y^{n+1}(\mathfrak{c}) = x^n(\mathfrak{c})$ and $wl^{n+1} = (1-\alpha)\lambda_{\mathfrak{j}^{n+1}(\mathfrak{c})} y^{n+1}(\mathfrak{c})$, we have $wl^{n+1}(\mathfrak{c}) = \alpha w l^n(\mathfrak{c})$. Chain feasibility requires that

$$y^0(\mathfrak{c}) \leq q(\mathfrak{c}) \sum_{k=0}^{\infty} \left(\frac{1}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}} l^n(\mathfrak{c})^{1-\alpha} \right)^{\alpha^n}$$

Using $l^n(\mathfrak{c}) = \alpha^n l^0(\mathfrak{c}) = \alpha^n (1-\alpha)\frac{\lambda(\mathfrak{c})}{w} y_j^0(\mathfrak{c})$, this is equivalent to

$$y^0(\mathfrak{c}) \leq q(\mathfrak{c}) \sum_{k=0}^{\infty} \left[\frac{1}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}} \left(\alpha^n (1-\alpha) \frac{y_j^0(\mathfrak{c})}{q(\mathfrak{c})} \right)^{1-\alpha} \right]^{\alpha^n} = y_j^0(\mathfrak{c})$$

Thus the allocation is chain feasible. ■

Lemma 9. *Given Φ, if $\{\hat{q}_j\}_{j\in[0,1]}$ satisfy $\hat{q}_j = \max_{\phi\in U_j} z(\phi)\hat{q}^{\alpha}_{s(\phi)}$ and $\hat{q}_j \leq sup_{\mathfrak{c}\in\mathfrak{C}_j} q(\mathfrak{c})$, then there exists a contracting arrangement that generates a feasible allocation with $q_j = \hat{q}_j$.*

Proof. For each j, if U_j is non-empty let $\phi^*_j \in \arg\max_{\phi\in U_j} z(\phi)\hat{q}^{\alpha}_{s(\phi)}$ (if there are multiple, pick one). Similarly, if $\mathfrak{C}_j$ is on-empty let $\mathfrak{c}^*_j$ be the chain generated by the $\{\phi^*_j\}$ to produce good j. We will construct an allocation by building the supply chain representation. Let $\hat{Q} = \int_0^1 \left(\hat{q}_j^{\varepsilon-1} dj\right)^{\frac{1}{\varepsilon-1}}$. For each j, let $y^0(\mathfrak{c}^*_j) = \hat{q}^{\varepsilon}_j \hat{Q}^{1-\varepsilon} L$. Then iteratively let $\frac{w}{\hat{q}_{j^n(\mathfrak{c})}} y^n(\mathfrak{c}) = \frac{wl^n(\mathfrak{c})}{1-\alpha} = \frac{\frac{w}{\hat{q}_{j^{n+1}(\mathfrak{c})}} x^n(\mathfrak{c})}{\alpha}$ and $x^n(\mathfrak{c}) = y^{n+1}(\mathfrak{c})$. Then each $x(\phi)$ is found by simply summing over the supply chains that pass through ϕ. The contracting arrangement with $\{x(\phi)\}$ (and any payments $\{T(\phi)\}$ thus generates an allocation with $q_j = \hat{q}_j$ that, by the previous lemma, is feasible. ■

A.4 Payoffs

To simplify further derivations, we can separate the payment for any technique into two parts. Given the arrangement and individual choices, we can define

$$\begin{aligned}
\tau(\phi) &\equiv T(\phi) - \lambda_{s(\phi)} x(\phi) \\
\pi^0_j &\equiv (p^0_j - \lambda_j) y^0_j = \frac{P^0 Y^0}{\varepsilon} \left(\frac{q_j}{Q}\right)^{\varepsilon-1}
\end{aligned}$$

$\tau(\phi)$ is the value of the payment above the value of the intermediate inputs when those inputs are priced at the supplier's marginal cost. π^0_j is the profit from sales of good j to the household when good j is valued at j's marginal cost.

Lemma 10. *In any pairwise stable arrangement, entrepreneur j's profit equals*

$$\pi_j = \pi^0_j + \sum_{\phi\in D_j} \tau(\phi) - \sum_{\phi\in U_j} \tau(\phi)$$

Proof. j's profit can be written as

$$\begin{aligned}
\pi_j &= (p^0_j - \lambda_j) y^0_j + \lambda_j y^0_j - w l_j + \sum_{\phi\in D_j} T(\phi) - \sum_{\phi\in U_j} T(\phi) \\
&= (p^0_j - \lambda_j) y^0_j + \lambda_j \left(y_j - \sum_{\phi\in D_j} x(\phi) \right) - w l_j + \sum_{\phi\in D_j} T(\phi) - \sum_{\phi\in U_j} T(\phi) \\
&= \pi^0_j + \lambda_j y_j - w l_j - \sum_{\phi\in U_j} x(\phi) + \sum_{\phi\in D_j} \tau(\phi) - \sum_{\phi\in U_j} \tau(\phi)
\end{aligned}$$

The conclusion follows from (19) and (20). ■

Recall that J^* is the set of entrepreneurs with acyclic networks and Φ^* is the set of techniques for which the buyer and supplier are members of J^*.

Proposition 2 (1 of 3)**.** *In any countably stable equilibrium, for any $\phi \in \Phi^*$, $\tau(\phi) \leq \sum_{j\in\mathcal{B}(\phi)} \left(\pi_j^0 - \pi_{j\setminus\phi}^0\right)$.*

Proof. Toward a contradiction, suppose there is a ϕ such that $\tau(\phi) > \sum_{j\in\mathcal{B}(\phi)} \left(\pi_j^0 - \pi_{j\setminus\phi}^0\right) + \eta$ with $\eta > 0$. Then there is a profitable deviation in which $b(\phi)$ drops technique ϕ and all entrepreneurs downstream from $s(\phi)$ increase production within their best alternative supply chain that does not pass through ϕ. The set $\mathcal{B}(\phi)$—entrepreneurs with chains that go through ϕ—is countable. Label these entrepreneurs by the natural number k. Each such entrepreneur has a supply chain that does not pass through ϕ that would deliver efficiency marginal cost $\tilde{\lambda}_{j^k}$ which is arbitrarily close to $\lambda_{j^k\setminus\phi}$, that satisfies $\pi_{j^k\setminus\phi}^0 - \frac{1}{\varepsilon} PY \left(\frac{w/\tilde{\lambda}_{j^k}}{Q}\right)^{\varepsilon-1} \leq \frac{\eta}{2^k}$. In the deviation, each entrepreneur in $\mathcal{B}(\phi)$ reduces production using the chain the passes through ϕ to zero, and increases production in their alternative chain in order to generate profit from sales to the household of $\frac{1}{\varepsilon} PY \left(\frac{w/\tilde{\lambda}_{j^k}}{Q}\right)^{\varepsilon-1}$. The deviation within each chain is the same as described in the proof that $q_j = \sup_{\mathfrak{c}\in\mathfrak{C}_j} q(\mathfrak{c})$. Among entrepreneurs in the deviation, the change in profit includes the recovery of the payment $\tau(\phi)$ minus the loss in profit from sales to the household is bounded above by

$$\begin{aligned}
& \tau(\phi) - \sum_{k=0}^{\infty} \left\{ \pi_{j^k}^0 - \frac{1}{\varepsilon} PY \left(\frac{w/\tilde{\lambda}_{j^k}}{Q}\right)^{\varepsilon-1} \right\} \\
= \; & \tau(\phi) - \sum_{k=1}^{\infty} \left(\pi_{j^k}^0 - \pi_{j^k\setminus\phi}^0\right) - \sum_{k=1}^{\infty} \left(\pi_{j^k\setminus\phi}^0 - \frac{1}{\varepsilon} PY \left(\frac{w/\tilde{\lambda}_{j^k}}{Q}\right)^{\varepsilon-1}\right) \\
\geq \; & \tau(\phi) - \sum_{k=1}^{\infty} \left(\pi_{j^k}^0 - \pi_{j^k\setminus\phi}^0\right) - \sum_{k=1}^{\infty} \frac{\eta}{2^k} \\
> \; & 0
\end{aligned}$$

The entire deviation involves a countable set of entrepreneurs because a countable union of countable sets is countable. ■

Proposition 2 (2 of 3)**.** *In any countably-stable equilibrium, for any $\phi \in \Phi^*$, $\tau(\phi) \geq 0$.*

Proof. If $T(\phi) < \lambda_{s(\phi)} x(\phi)$ then the supplier would gain by dropping the contract and reducing production throughout its supply chain. The cost savings to the chain would be $\lambda_{s(\phi)} x(\phi)$ which is larger than the payment $T(\phi)$ from $b(\phi)$. ■

A.5 Existence and Bargaining Power

For any coalition J, let $\mathcal{U}(J)$ and $\mathcal{D}(J)$ be the set of techniques that are respectively directly upstream and directly downstream from members of J, so that $\mathcal{U}(J) = \{\phi | b(\phi) \in J, s(\phi) \notin J\}$ and $\mathcal{D}(J) = \{\phi | s(\phi) \in J, b(\phi) \notin J\}$. The sum of the payoffs to members of J is

$$\sum_{j \in J} \pi_j^0 - \sum_{\phi \in \mathcal{U}(J)} \tau(\phi) + \sum_{\phi \in \mathcal{D}(J)} \tau(\phi) \tag{22}$$

If a coalition J deviates, there is a subset of techniques in $\mathcal{U}(J)$ that are dropped. Let $\mathcal{U}^-(J)$ be the entrepreneurs that are the suppliers of those techniques. We also define $\lambda_{j \setminus \mathcal{U}^-(J)}$ and $\pi^0_{j \setminus \mathcal{U}^-(J)}$ to be what j's marginal cost and profit from sales to the household would be if it were unable to use chains that passed through any of those entrepreneurs in $\mathcal{U}^-(J)$.

Lemma 11. *A contracting arrangement that generates a feasible allocation and satisfies* $q_j = \sup_{\mathfrak{c} \in \mathfrak{C}_j} q(\mathfrak{c})$ *for each* j *is countably stable if and only if for each coalition* J *and each subset of upstream techniques* $\mathcal{U}^-(J)$*, the following equation holds:*

$$\sum_{j \in J} \pi_j^0 - \pi^0_{j \setminus \mathcal{U}^-(J)} + \sum_{\phi \in \mathcal{D}(J)} \min\left\{\tau(\phi), -\left(\lambda_{s(\phi)} - \lambda_{s(\phi) \setminus \mathcal{U}^-(J)}\right) x(\phi)\right\} \geq \sum_{\phi \in \mathcal{U}^-(J)} \tau(\phi) \tag{23}$$

Proof. Consider an arrangement such that the resulting allocation is feasible and $q_j = \sup_{\mathfrak{c} \in \mathfrak{C}_j} q(\mathfrak{c})$ for all j.

Suppose first that the resulting allocation is countably stable. Consider a coalition J and suppose there is a deviation where $\mathcal{U}^-(J)$ are the upstream contracts that are dropped. For any technique downstream from the coalition ($\phi \in \mathcal{D}(J)$), the coalition could either drop the contract or continue to supply those inputs. Following the deviation, the sum of the payoffs to the members of J is no better than

$$\sum_{j \in J} \pi^0_{j \setminus \mathcal{U}^-(J)} - \sum_{\phi \in \mathcal{U}(J) \setminus \mathcal{U}^-(J)} \tau(\phi) + \sum_{\phi \in \mathcal{D}(J)} \max\left\{0, T(\phi) - \lambda_{s(\phi) \setminus \mathcal{U}^-(J)} x(\phi)\right\} \tag{24}$$

Further, there is a larger coalition (the union of the coalition J and those entrepreneurs involved in the alternative supply chains) with a deviation that would attain payoffs (24) for those in J and leave the other entrepreneurs in that larger deviation no worse off. Countable stability means the (23) is weakly greater than (24), which implies (23).

Next suppose (23) holds for all J and $\mathcal{U}^-(J)$. Then following any deviation by any coalition J that drops suppliers $\mathcal{U}^-(J)$, the payoff following the deviation is no greater than (24). Since (23) implies that this is weakly less than (24), the deviation cannot dominate. ■

Lemma 12. *In any contracting arrangement that generates a feasible allocation that satisfies $q_j = \sup_{\mathfrak{c}\in\mathfrak{C}_j} q(\mathfrak{c})$ for all j, for any $\phi, \phi' \in \Phi^*$ such that ϕ' is downstream of ϕ,*

$$-\left[\lambda_{s(\phi')} - \lambda_{s(\phi')\setminus\phi}\right] x(\phi') \geq \sum_{j\in\mathcal{B}(\phi')} \left(\pi_j^0 - \pi_{j\setminus s(\phi)}^0\right) \tag{25}$$

Proof. Suppose first that $\lambda_{s(\phi')} = \lambda_{s(\phi')\setminus\phi}$. Then it must be that for each $j \in \mathcal{B}(\phi')$ that $\pi_j^0 = \pi_{j\setminus s(\phi)}^0$, which implies (25).

Suppose instead that $\lambda_{s(\phi')} < \lambda_{s(\phi')\setminus\phi}$. Consider a supply chain representation of the allocation. For any $j \in \mathcal{B}(\phi')$, there is the subset of chains $\mathfrak{C}_j$ that pass through ϕ' which we label let $\mathfrak{C}_j(\phi')$, and a k_j such that ϕ' is the k_j-th technique in every $\mathfrak{c} \in \mathfrak{C}_j(\phi')$. Then

$$x(\phi') = \sum_{j\in\mathcal{B}(\phi')} \sum_{\mathfrak{c}\in\mathfrak{C}_j(\phi')} x^{k_j}(\mathfrak{c}) \tag{26}$$

If $\sum_{\mathfrak{c}\in\mathfrak{C}_j(\phi')} y^0(\mathfrak{c}) < y_j^0$, then it must be that $\pi_j^0 = \pi_{j\setminus s(\phi')}^0$, which implies $\pi_j^0 = \pi_{j\setminus s(\phi)}^0$. If, on the other hand, $\sum_{\mathfrak{c}\in\mathfrak{C}_j(\phi')} y^0(\mathfrak{c}) = y_j^0$, then

$$\pi_j^0 = p_j^0 y_j^0 - \sum_{\mathfrak{c}\in\mathfrak{C}_j(\phi')} \left[w \sum_{n=0}^{k_j} l^n(\mathfrak{c}) - \lambda_{s(\phi')} x^{k_j}(\mathfrak{c}) \right]$$

We can also define $\tilde{\pi}_j^0$ to be

$$\tilde{\pi}_j^0 \equiv p_j^0 y_j^0 - \sum_{\mathfrak{c}\in\mathfrak{C}_j(\phi')} \left[w \sum_{n=0}^{k_j} l^n(\mathfrak{c}) - \lambda_{s(\phi')\setminus\phi} x^{k_j}(\mathfrak{c}) \right]$$

Note first that $\pi_j^0 - \tilde{\pi}_j^0 = \sum_{\mathfrak{c}\in\mathfrak{C}_j(\phi')} - \left[\lambda_s - \lambda_{s(\phi')\setminus\phi}\right] x^n(\mathfrak{c}_j)$. Note second that $\tilde{\pi}_j^0 \leq \pi_{j\setminus s(\phi)}^0$, because if j could not use chains that passed through ϕ it could reoptimize its use of labor or choose alternative chains. Together these imply that

$$\pi_j^0 - \pi_{j\setminus s(\phi)}^0 \leq \sum_{\mathfrak{c}\in\mathfrak{C}_j(\phi')} - \left[\lambda_s - \lambda_{s(\phi')\setminus\phi}\right] x^{k_j}(\mathfrak{c}_j)$$

Summing over $j \in \mathcal{B}(\phi')$ and using (26) gives

$$\begin{aligned} -\left[\lambda_{s(\phi')} - \lambda_{s(\phi')\backslash\phi}\right] x(\phi') &= \sum_{j \in \mathcal{B}(\phi')} \sum_{\mathfrak{c} \in \mathfrak{C}_j(\phi')} -\left[\lambda_{s(\phi')} - \lambda_{s(\phi')\backslash\phi}\right] x^n(\mathfrak{c}) \\ &\geq \sum_{j \in \mathcal{B}(\phi')} \pi_j^0 - \pi_{j\backslash s(\phi)}^0 \end{aligned}$$

■

Lemma 13. *Suppose there is a coalition in J^* with a dominating deviation. Then there is dominating deviation in which at most a single technique is dropped.*

Proof. Suppose first that there is a technique $\phi \in \Phi^*$ such that $\tau(\phi) < 0$. Then by the argument of Proposition 2(2), there is a dominating deviation in which no suppliers are dropped. Suppose instead that $\tau(\phi) \geq 0$ for all $\phi \in \Phi^*$. Toward a contradiction, suppose that there is no dominating deviation in which a single technique is dropped. We will show that this implies there is no dominating deviation in which multiple suppliers are dropped. If there are no deviations in which a single contract is dropped, it must be that for each $\tilde{\phi} \in \mathcal{U}^-(J)$

$$\sum_{j \in J} \pi_j^0 - \pi_{j\backslash\tilde{\phi}}^0 + \sum_{\phi \in \mathcal{D}(J)} \min\left\{\tau(\phi), -\left(\lambda_{s(\phi)} - \lambda_{s(\phi)\backslash\tilde{\phi}}\right) x(\phi)\right\} \geq \tau\left(\tilde{\phi}\right)$$

Summing across $\tilde{\phi} \in \mathcal{U}^-(J)$,

$$\sum_{\tilde{\phi} \in \mathcal{U}^-(J)} \sum_{j \in J} \pi_j^0 - \pi_{j\backslash\tilde{\phi}}^0 + \sum_{\tilde{\phi} \in \mathcal{U}^-(J)} \sum_{\phi \in \mathcal{D}(J)} \min\left\{\tau(\phi), -\left(\lambda_{s(\phi)} - \lambda_{s(\phi)\backslash\tilde{\phi}}\right) x(\phi)\right\} \geq \sum_{\tilde{\phi} \in \mathcal{U}^-(J)} \tau(\phi) \quad (27)$$

Consider any $j \in J$. Define $\hat{\phi}_j \in \arg\max_{\tilde{\phi} \in \mathcal{U}^-(J)} \lambda_{j\backslash\tilde{\phi}}$. That is, of all the techniques in $\mathcal{U}^-(J)$, if j could not use chains that passed through $\hat{\phi}_j$ its marginal cost would rise the most (if there are multiple such techniques, select one at random). Note that since $j \in J^*$ and J is connected set, any supply chain that goes through $\hat{\phi}_j$ does not go through any other technique in $\mathcal{U}^-(J)$. Thus while $\pi_j^0 > \pi_{j\backslash\hat{\phi}_j}^0$ and $\lambda_{j\backslash\hat{\phi}_j} \geq \lambda_j$, for any other technique $\phi' \in \mathcal{U}^-(J) \setminus \{\hat{\phi}_j\}$, $\pi_j^0 = \pi_{j\backslash\phi'}^0$ and $\lambda_{j\backslash\phi'} = \lambda_j$. Therefore summing across all techniques in $\mathcal{U}^-(J)$ gives two relationships. For each $j \in J$:

$$\sum_{\tilde{\phi} \in \mathcal{U}^-(J)} \pi_j^0 - \pi_{j\backslash\tilde{\phi}}^0 = \pi_j^0 - \pi_{j\backslash\hat{\phi}_j}^0 \leq \pi_j^0 - \pi_{j\backslash\mathcal{U}^-(J)}^0 \quad (28)$$

and for each $\phi \in \mathcal{D}(J)$

$$\begin{aligned}\sum_{\tilde{\phi}\in\mathcal{U}^-(J)} \min\left\{\tau(\phi), -\left(\lambda_{s(\phi)} - \lambda_{s(\phi)\backslash\tilde{\phi}}\right) x(\phi)\right\} &= \min\left\{\tau(\phi), -\left(\lambda_{s(\phi)} - \lambda_{s(\phi)\backslash\hat{\phi}_{s(\phi)}}\right) x(\phi)\right\} \\ &\leq \min\left\{\tau(\phi), -\left(\lambda_{s(\phi)} - \lambda_{s(\phi)\backslash\mathcal{U}^-(J)}\right) x(\phi)\right\} \quad (29)\end{aligned}$$

where, in each case, the inequality follows because $\pi^0_{j\backslash\hat{\phi}_j} \geq \pi^0_{j\backslash\mathcal{U}^-(J)}$ and $\lambda_{s(\phi)\backslash\hat{\phi}_{s(\phi)}} \leq \lambda_{s(\phi)\backslash\mathcal{U}^-(J)}$ respectively; the latter is more constrained. Summing (28) across $j \in J$ and (29) across $\phi \in \mathcal{D}(J)$, and then reversing the order of each summation gives

$$\sum_{\tilde{\phi}\in\mathcal{U}^-(J)} \sum_{j\in J} \pi^0_j - \pi^0_{j\backslash\tilde{\phi}} \leq \sum_{j\in J} \pi^0_j - \pi^0_{j\backslash\mathcal{U}^-(J)} \quad (30)$$

$$\sum_{\tilde{\phi}\in\mathcal{U}^-(J)} \sum_{\phi\in\mathcal{D}(J)} \min\left\{\tau(\phi), -\left(\lambda_{s(\phi)} - \lambda_{s(\phi)\backslash\tilde{\phi}}\right) x(\phi)\right\} \leq \sum_{\phi\in D(J)} \min\left\{\tau(\phi), -\left(\lambda_{s(\phi)} - \lambda_{s(\phi)\backslash U^-(J)}\right) x(\phi)\right\} \quad (31)$$

These and (27) imply (23) so that there is no dominating deviation for J. ■

Proposition 2 (3 of 3). *For any $\beta \in [0,1]$, there exists a countably-stable equilibrium in which $\tau(\phi) = \beta\mathcal{S}(\phi), \forall\phi \in \Phi^*$ and $\tau(\phi) = 0, \forall\phi \notin \Phi^*$.*

Proof. For each coalition J such that no member of J is in J^*, $\tau(\phi) = 0$ implies that (23) holds. Consider now a deviation by a coalition $J \subset J^*$ in which at most one technique is dropped. Divide J into two disjoint groups: J_1 are those that are downstream from the technique that is dropped, of the empty set if no technique is dropped, and J_2 are those that are not downstream from the technique that is dropped (or the empty set if all of $J_1 = J$). For Similarly, divide $\mathcal{D}(J)$ into $\mathcal{D}_1$ and $\mathcal{D}_2$, those techniques in $\mathcal{D}(J)$ downstream from entrepreneurs in J_1 and J_2 respectively.

For each technique $\phi' \in \mathcal{D}_1$, we have from (25) that if ϕ is the technique that is dropped,

$$-\left[\lambda_{s(\phi')} - \lambda_{s(\phi')\backslash\phi}\right] x(\phi') \geq \sum_{j\in\mathcal{B}(\phi')} \pi^0_j - \pi^0_{j\backslash s(\phi)} \geq \beta \sum_{j\in\mathcal{B}(\phi')} \pi^0_j - \pi^0_{j\backslash s(\phi)}$$

We also know that for each $\phi' \in \mathcal{D}_1$, we have that

$$\beta \sum_{j\in\mathcal{B}(\phi')} \pi^0_j - \pi^0_{j\backslash s(\phi)} \leq \beta \sum_{j\in\mathcal{B}(\phi')} \pi^0_j - \pi^0_{j\backslash s(\phi')}$$

Together, these imply that for each $\phi' \in \mathcal{D}_1$,

$$\min\left\{\tau(\phi'), -\left[\lambda_{s(\phi')} - \lambda_{s(\phi')\backslash\phi}\right] x(\phi')\right\} \geq \beta \sum_{j\in\mathcal{B}(\phi')} \pi_j^0 - \pi_{j\backslash s(\phi)}^0 \tag{32}$$

(32) also holds for each $\phi' \in \mathcal{D}_2$ because $\lambda_{s(\phi')} = \lambda_{s(\phi')}$ and $\pi_j^0 = \pi_{j\backslash s(\phi)}^0$. Putting these pieces together, we have that

$$\begin{aligned}\tau(\phi) &= \beta \sum_{j\in J} \pi_j^0 - \pi_{j\backslash\phi}^0 + \beta \sum_{\phi'\in\mathcal{D}(J)} \sum_{j\in\mathcal{B}(\phi')} \pi_j^0 - \pi_{j\backslash\phi}^0 \\ &\leq \sum_{j\in J} \pi_j^0 - \pi_{j\backslash\phi}^0 + \sum_{\phi'\in\mathcal{D}(J)} \min\left\{\tau(\phi'), -\left[\lambda_{s(\phi')} - \lambda_{s(\phi)\backslash\phi}\right] x(\phi')\right\}\end{aligned}$$

This is equivalent to (23) for the case in which $\mathcal{U}(J)$ is a singleton, so there is no coalition $J \in J^*$ with a dominating deviation. ■

B Proof of Proposition 3

The strategy begins with defining a sequence of random variables $\{X_N\}_{N\in\mathbb{N}}$ with the property that the maximum feasible efficiency of an entrepreneur 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^* in $L^{\varepsilon-1}$. Next we show that the CDF of X^* is the unique fixed point of T in $\mathcal{F}$, a subset of $\bar{\mathcal{F}}$ (and that such a fixed point exists). Letting F^* be this fixed point, the law of large numbers implies the CDF of the cross-sectional distribution of efficiencies is F^* and that aggregate productivity is $\|X^*\|_{\varepsilon-1}$.

B.1 Existence of a Fixed Point of Equation (11)

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 $\rho = e^{-M(1-\rho)}$. In the definition of $\bar{f}$, let $\beta > \varepsilon - 1$ be such that $\lim_{z\to\infty} z^\beta[1 - H(z)] = 0$. Then there exists a $z_2 > 1$ such that $z > z_2$ implies $z^\beta[1 - H(z)] < (1 - \alpha)$. With this, let q_2 be a number large enough so that $q_2^{(1-\alpha)\beta} > M\left(z_2^\beta + 1\right)$ and $q_2^{1-\alpha} > z_2$. In the definition of $\underline{f}$, $q_0 = z_0^{\frac{1}{1-\alpha}}$.

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

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

$$\underline{f}(q) \equiv \begin{cases} \rho, & q < q_0 \\ 1, & q \geq q_0 \end{cases}$$

On $\bar{\mathcal{F}}$, the set of right continuous, weakly increasing functions $f : \mathbb{R}^+ \to [0,1]$, consider the partial order 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 14. *$T\underline{f} \preceq \underline{f}$ and $\bar{f} \preceq T\bar{f}$*

Proof. We first show $T\underline{f} \preceq \underline{f}$. For $q \geq q_0$, $T\underline{f}(q) \leq 1 = \underline{f}(q)$. For $q < q_0$:

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

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

$$\int_{z_0}^{q/q_2^\alpha} z^{\frac{\beta}{\alpha}} dH(z) + (q/q_2^\alpha)^{\frac{\beta}{\alpha}} \left[1 - H\left(q/q_2^\alpha\right)\right] < \left(z_2^\beta + 1\right) (q/q_2^\alpha)^{\frac{1-\alpha}{\alpha}\beta} \tag{33}$$

To see this note we can integrate by parts to get

$$\int_{z_2}^{q/q_2^\alpha} z^{\beta/\alpha} dH(z) = \left[1 - H\left(z_2\right)\right] z_2^{\beta/\alpha} - (q/q_2^\alpha)^{\beta/\alpha} \left[1 - H\left(q/q_2^\alpha\right)\right] + \int_{z_2}^{q/q_2^\alpha} \frac{\beta}{\alpha} z^{\beta/\alpha - 1} \left[1 - H(z)\right] dz$$

Rearranging this gives

$$H\left(z_2\right) z_2^{\beta/\alpha} + \int_{z_2}^{q/q_2^\alpha} z^{\beta/\alpha} dH(z) + (q/q_2^\alpha)^{\beta/\alpha} \left[1 - H\left(q/q_2^\alpha\right)\right] = z_2^{\beta/\alpha} + \frac{\beta}{\alpha} \int_{z_2}^{q/q_2^\alpha} z^{\beta/\alpha - 1} \left[1 - H(z)\right] dz$$

Since $q/q_2^\alpha > z_2$, equation (33) follows from this and three inequalities: (i) $\int_{z_0}^{z_2} z^{\beta/\alpha} dH(z) \leq H\left(z_2\right) z_2^{\beta/\alpha}$; (ii) $z_2^{\beta/\alpha} \leq z_2^\beta \left(q/q_2^\alpha\right)^{\beta/\alpha - \beta}$; and (iii) $\int_{z_2}^{q/q_2^\alpha} z^{\beta/\alpha - 1} \left[1 - H(z)\right] dz \leq \int_0^{q/q_2^\alpha} z^{\beta/\alpha - 1} \left[(1-\alpha) z^{-\beta}\right] dz$.

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

$$\begin{aligned}
\frac{1-T\bar{f}(q)}{1-\rho} &\leq M\int_{z_0}^{\infty}\frac{1-\bar{f}\left((q/z)^{1/\alpha}\right)}{1-\rho}dH(z)\\
&= M\int_{z_0}^{q/q_2^{\alpha}}\left(\frac{(q/z)^{1/\alpha}}{q_2}\right)^{-\beta}dH(z)+M\left[1-H\left(q/q_2^{\alpha}\right)\right]\\
&= \left(\frac{q}{q_2}\right)^{-\beta}\frac{M\left(z_2^{\beta}+1\right)}{q_2^{(1-\alpha)\beta}}\left\{\frac{\int_{z_0}^{q/q_2^{\alpha}}z^{\frac{\beta}{\alpha}}dH(z)+(q/q_2^{\alpha})^{\frac{\beta}{\alpha}}\left[1-H\left(q/q_2^{\alpha}\right)\right]}{\left(z_2^{\beta}+1\right)(q/q_2^{\alpha})^{\frac{1-\alpha}{\alpha}\beta}}\right\}\\
&\leq \left(\frac{q}{q_2}\right)^{-\beta}\\
&= \frac{1-\bar{f}(q)}{1-\rho}
\end{aligned}$$

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

Lemma 15. *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. The operator T is order preserving, and $\mathcal{F}$ is a complete lattice. By the Tarski fixed point theorem, the set of fixed points of T in $\mathcal{F}$ is also a complete lattice, and hence has a least and a greatest fixed point given by $\lim_{N\to\infty}T^N\bar{f}$ and $\lim_{N\to\infty}T^N\underline{f}$ respectively. ∎

B.2 Existence of a Limit

For and chain $\mathfrak{c}\in\mathfrak{C}_j$, define $\mathfrak{q}_N(\mathfrak{c})\equiv\prod_{n=0}^{N-1}z^n(\mathfrak{c})^{\alpha^n}$ for $N\geq 1$ and $\mathfrak{q}_0(\mathfrak{c})\equiv 1$. For any chain in $\mathfrak{C}_j$, a subchain of length N is the segment of techniques of length N that is most downstream. Let $\mathfrak{C}_{j,N}$ be the set of *distinct* subchains of length N.

With these, we will define three sequences of random variables for each entrepreneur, $\{X_{j,N},\bar{Y}_{j,N},\underline{Y}_{j,N}\}$ so that their respective CDFs are T^Nf^1, $T^N\bar{f}$, and $T^N\underline{f}$. The construction is guided by the following lemma.

Lemma 16. *Given Φ, define the random variables $\{\hat{q}(\mathfrak{c})\}_{\forall\mathfrak{c}\in\bigcup_{N=0}^{\infty}\mathfrak{C}_{j,N},\forall j}$ to be IID random variables with CDF $\frac{\hat{f}-\rho}{1-\rho}$. For each j, N and for each $\mathfrak{c}\in\mathfrak{C}_{j,N}$, let $\hat{\mathfrak{q}}_N(\mathfrak{c})=\mathfrak{q}_N(\mathfrak{c})\hat{q}(\mathfrak{c})^{\alpha^N}$. Let $\hat{Y}_{j,N}=\max_{\mathfrak{c}\in\mathfrak{C}_{j,N}}\hat{\mathfrak{q}}_N(\mathfrak{c})$. Then the CDF of $\hat{Y}_{j,N}$ is $T^N\hat{f}$.*

Proof. We proceed by induction. For each j, let $U_j^\infty \subseteq U_j$ be the techniques that are the most downstream techniques of at least one supply chain $\mathfrak{c} \in \mathfrak{C}_j$. If $\phi \notin U_j^\infty$ then it is infeasible for j to acquire the inputs to produce using ϕ.

Consider a single technique ϕ in U_j. A standard result from the theory of branching processes is that the probability $\phi \in U_j^\infty$ is $1-\rho$ (see Online Appendix B.4 for a derivation). In that case, the probability that $z(\phi)\hat{q}(\mathfrak{c})^\alpha \leq q$ is $\int \frac{f((q/z)^{1/\alpha})-\rho}{1-\rho} dH(z)$. Since $\hat{Y}_{j,1} = \max_{\mathfrak{c}\in\mathfrak{C}_{j,1}} \hat{\mathfrak{q}}_1(\mathfrak{c}) = \max_{\phi \in U_j^\infty} z(\phi)\hat{q}(\mathfrak{c})^\alpha$, summing over the possible realizations of U_j gives

$$
\begin{aligned}
\Pr\left(\hat{Y}_{j,1} \leq q\right) &= \sum_{k=0}^\infty \frac{e^{-M}M^k}{k!}\left[\Pr(\phi \notin U_j^\infty \text{ or } z(\phi)\hat{q}(\mathfrak{c})^\alpha)\right]^k \\
&= \sum_{k=0}^\infty \frac{e^{-M}M^k}{k!}\left[\rho + (1-\rho)\int \frac{f\left((q/z)^{1/\alpha}\right)-\rho}{1-\rho} dH(z)\right]^k \\
&= e^{-M\left[1-\int f\left((q/z)^{1/\alpha}\right)dH(z)\right]} = Tf
\end{aligned}
$$

Now suppose the CDF of $Y_{s(\phi),N}$ is $T^N f$. Using the logic behind equation (11), for any $\phi \in U_j$, $\Pr\left(z(\phi)\hat{Y}_{s(\phi),N}^\alpha \leq q\right) = \int T^N f\left((q/z)^{1/\alpha}\right)dH(z)$, so that integrating over realizations of U_j gives

$$
\Pr\left(\hat{Y}_{j,N+1} \leq q\right) = \Pr\left(\max_{\phi\in U_j} z(\phi)\hat{Y}_{s(\phi),N}^\alpha \leq q\right) = e^{-M\left[1-\int T^N f\left((q/z)^{1/\alpha}\right)dH(z)\right]} = T^{N+1}f
$$

■

If $\mathfrak{C}_j$ is empty, then define $X_{j,n} = \bar{Y}_{j,N} = \underline{Y}_{j,N} = 0$ for all $N \geq 0$. If $\mathfrak{C}_j$ is non-empty, we define $X_{j,N} = \sup_{\mathfrak{c}\in\mathfrak{C}_j} \mathfrak{q}_N(\mathfrak{c})$. Roughly, the remainder of this subsection shows that $X_{j,N}$ converges to q_j. Since $q_j = \sup_{\mathfrak{c}\in\mathfrak{C}_j} \lim_{N\to\infty} \mathfrak{q}_N(\mathfrak{c})$, we are essentially proving that the limit can be passed through the sup.

One consequence of Lemma 16 is that the CDF of $X_{j,N}$ is $T^N f^1$ (the variables $\hat{q}(\mathfrak{c})$ are all degenerate and equal to 1). To construct $\bar{Y}_{j,N}$, given a realization of Φ, let $\{\bar{q}(\mathfrak{c})\}_{\forall \mathfrak{c}\in\bigcup_{N=0}^\infty \mathfrak{C}_{j,N},\forall j}$ be IID random variables, each with CDF $\frac{\bar{f}-\rho}{1-\rho}$. With this we define $\bar{\mathfrak{q}}_N(\mathfrak{c}) \equiv \mathfrak{q}_N(\mathfrak{c})\bar{q}(\mathfrak{c})^{\alpha^N}$ and $\underline{\mathfrak{q}}_N(\mathfrak{c}) \equiv \mathfrak{q}_N(\mathfrak{c})q_0^{\alpha^N}$ (q_0 is the same as a random variable with CDF $\frac{f-\rho}{1-\rho}$). Lastly, for $N \geq 1$, let $\bar{Y}_{j,N} \equiv \max_{\mathfrak{c}\in\mathfrak{C}_{j,N}} \bar{\mathfrak{q}}_N(\mathfrak{c})$ and $\underline{Y}_{j,N} \equiv \max_{\mathfrak{c}\in\mathfrak{C}_{j,N}} \underline{\mathfrak{q}}_N(\mathfrak{c})$. Also let $X_{j,0} = 1$, $\underline{Y}_{j,0} = q_0$, and $\bar{Y}_{j,0}$ to have CDF $\frac{\bar{f}-\rho}{1-\rho}$.

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

Lemma 17. *For each $N \geq 0$, X_N, $\bar{Y}_N$, and $\underline{Y}_N$ are uniformly integrable in $L^{\varepsilon-1}$.*

Proof. First, recall that $\bar{Y}_0$ is defined so that its CDF is $\bar{f}$. Since 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 18. *There exists a random variable X^* such that X_N converges to X^* almost surely and in $L^{\varepsilon-1}$.*

Proof. Let $P_N \equiv \frac{X_N}{\prod_{n=0}^{N}\mu_n}$ where $\mu_n \equiv M\int_{z_0}^{\infty} w^{\alpha^n}\rho^{1-H(w)}dH(w)$. We first show that $\{P_N\}$ is a submartingale with respect to $\{\mathfrak{C}_N\}$.

If $\mathfrak{C}$ is empty then $\mathbb{E}\left[P_N|\mathfrak{C}_{N-1}\right] = P_{N-1} = 0$. Otherwise, define a set $\mathcal{D}_N$ as follows: Let $\mathfrak{c}_N^* \in \arg\max_{\mathfrak{c}\in\mathfrak{C}_N} \mathfrak{q}_N(\mathfrak{c})$ so that $X_N = \mathfrak{q}_N\left(\mathfrak{c}_N^*\right)$. Let $\mathcal{D}_N \subseteq \mathfrak{C}_N$ be the set of chains in $\mathfrak{C}_N$ for which the first $N-1$ links are $\mathfrak{c}_{N-1}^*$. In other words, all chains in $\mathcal{D}_N$ are of the form $\mathfrak{c}_{N-1}^*\phi$ for some ϕ.

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

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

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

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

Given X_{N-1}, if $\mathcal{D}_N$ consists of k chains, the probability that $D_N < x$ is

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

With this, the CDF of D_N, given $\mathfrak{C}_{N-1}$, is

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

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

$$\mathbb{E}\left[D_N|\mathfrak{C}_{N-1}\right] = X_{N-1}\int_{z_0}^{\infty} w^{\alpha^N}\log\rho^{-1}\frac{\rho^{1-H(w)}}{1-\rho}dH(w) = \mu_N X_{N-1}$$

Putting this together, we have

$$\mathbb{E}\left[P_N|\mathfrak{C}_{N-1}\right] = \frac{1}{\prod_{n=0}^{N}\mu_n}\mathbb{E}\left[X_N|\mathfrak{C}_{N-1}\right] \geq \frac{1}{\prod_{n=0}^{N}\mu_n}\mathbb{E}\left[D_N|\mathfrak{C}_{N-1}\right] = \frac{1}{\prod_{n=0}^{N}\mu_n}\mu_N X_{N-1} = P_{N-1}$$

We next show that $\{P_N\}$ is uniformly integrable, i.e., that $\sup_N \mathbb{E}\left[P_N\right] < \infty$. Since $\sup_N \mathbb{E}\left[X_N\right] < \infty$, it suffices to show a uniform lower bound on $\left\{\prod_{n=0}^{N}\mu_n\right\}$. 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 $\{P_N\}_{N\in\mathbb{N}}$ is a uniformly integrable (in L^1) submartingale, so by the martingale convergence theorem, there exists a P such that P_N converges to P almost surely. By the continuous mapping theorem, there exists an X^* such that X_N converges to X^* 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^* in $L^{\varepsilon-1}$. ■

Lemma 19. *If $\mathfrak{C}$ is nonempty then with probability one, $X^* = \sup_{\mathfrak{c}\in\mathfrak{C}}\mathfrak{q}(\mathfrak{c})$*

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

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

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

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

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 $\sup_{\mathfrak{c}\in\mathfrak{C}} \mathfrak{q}(\mathfrak{c}) \geq X^*$ with probability one. ■

B.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^* in probability. Convergence in probability implies convergence in distribution, which gives two 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^*, 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^* is the CDF of X^*.

We first show that $\log \bar{Y}_N - \log \underline{Y}_N$ converges to zero in probability.

Lemma 20. *If $\mathfrak{C}$ is nonempty, then for any $\eta > 1$, $\lim_{N\to\infty} \Pr\left(\bar{Y}_N / \underline{Y}_N > \eta\right) = 0$.*

Proof. Let $\mathfrak{S}_{j,N}$ be the set of chain stubs: sequences of N techniques $\{\phi_n\}_{n=0}^{N-1}$ such that $s(\phi_n) = b(\phi_{n+1})$ and $b(\phi_0) = j$. Note that each $\mathfrak{c} \in \mathfrak{C}_{j,N}$ is a chain stub so that $\mathfrak{C}_{j,N} \subseteq \mathfrak{S}_{j,N}$, but not the other way around because each subchain $\mathfrak{c} \in \mathfrak{C}_{j,N}$ must satisfy the additional requirement that there is a chain in $\mathfrak{C}_j$ for which it is the N most downstream techniques.

Conditional on the set of supply chains Φ, we have that for each chain $\mathfrak{c}$ in $\mathfrak{C}_N$ that $bar\mathfrak{q}_N(\mathfrak{c})/\underline{\mathfrak{q}}_N(\mathfrak{c}) = \left(\frac{\bar{q}(\mathfrak{c})}{q_0}\right)^{\alpha^N}$. We therefore have:

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

If $|\mathfrak{C}_N|$ is the number of distinct subchains of length N, the probability that every chain $\mathfrak{c} \in \mathfrak{C}_N$

satisfies $\bar{\mathfrak{q}}_N(\mathfrak{c})/\underline{\mathfrak{q}}_N(\mathfrak{c}) \leq \eta$ is $\left(\frac{\bar{f}\left(\eta^{\alpha^{-N}} q_0\right)-\rho}{1-\rho}\right)^{|\mathfrak{C}_N|}$ so that

$$\Pr\left(\bar{Y}_N/\underline{Y}_N \leq \eta|\Phi\right) = \left(\frac{\bar{f}\left(\eta^{\alpha^{-N}} q_0\right)-\rho}{1-\rho}\right)^{|\mathfrak{C}_N|} \geq \left(\frac{\bar{f}\left(\eta^{\alpha^{-N}} q_0\right)-\rho}{1-\rho}\right)^{|\mathfrak{S}_N|}$$

A standard result from the theory of branching processes (see Online Appendix B.4 for a derivation) is that for any x, $\mathbb{E}\left[x^{|\mathfrak{S}_N|}\right] = \varphi^{(N)}(x)$ where $\varphi^{(N)}$ is the N-fold composition of $\varphi(x) \equiv e^{-M(1-x)}$, the probability generating function for $|\mathfrak{S}_1|$, and expectations are taken over realizations of Φ. This implies

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

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 showing $\lim_{N\to\infty} 1-\varphi^{(N)}\left(\frac{\bar{f}\left(\eta^{\alpha^{-N}} q_0\right)-\rho}{1-\rho}\right) = 0$.

To do this, we first show that for $x \in [0,1]$, $\frac{d}{dx}\varphi^{(N)}(x) \leq M^N$. To see this, note that φ is convex and $\varphi'(1) = M$, so that $\varphi'(x) \leq M$ 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

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

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

$$\varphi^{(N)}(x) = \varphi^{(N)}(1) - \int_x^1 \varphi^{(N)\prime}(w)dw \geq 1 - M^N \int_x^1 dw \geq 1 - M^N[1-x]$$

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

We now come to the main result.

Proposition 3. *There is a unique fixed point of T on $\mathcal{F}$, F^*. F^* is CDF of X^*. Aggregate productivity is $Q = \left(\int_0^\infty q^{\varepsilon-1} dF^*(q)\right)^{\frac{1}{\varepsilon-1}}$ with probability one.*

Proof. If $\mathfrak{C}$ is nonempty, 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 $X_N \xrightarrow{p} X^*$ implies that $\bar{Y}_N \xrightarrow{p} X^*$ and $\underline{Y}_N \xrightarrow{p} X^*$. If $\mathfrak{C}$ is empty, then $X^* = 0$, so that $\bar{Y}_N \xrightarrow{p} X^*$ and $\underline{Y}_N \xrightarrow{p} X^*$. Together, these imply that $\bar{Y}_N \xrightarrow{p} X^*$ and $\underline{Y}_N \xrightarrow{p} X^*$ unconditionally.

We first show that there is a unique fixed point, which is also the CDF of X^*. 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^*. Call this fixed point F^*.

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

Putting all of these pieces together, we have that the CDF of q_j is F^*. We next show that aggregate productivity is the $Q = \left(\int_0^\infty q^{\varepsilon-1} dF^*(q)\right)^{\frac{1}{\varepsilon-1}}$. 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 entrepreneurs, j and i. Since the set of entrepreneurs 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 that X^* is $L^{\varepsilon-1}$ integrable. Therefore we have that $Q = \left(\int_0^\infty q^{\varepsilon-1} dF^*(q)\right)^{\frac{1}{\varepsilon-1}}$ with probability one. ■

B.4 The Number of Supply Chains

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

Let $B_{j,N} \equiv |\mathfrak{S}_{j,N}|$ be the number of distinct chains stubs of length N. Let $p(k)$ be the probability that an entrepreneur has exactly k techniques, in this case equal to $\frac{e^{-M} M^k}{k!}$, and let $P_N(l,k)$ be the probability that, in total, l different entrepreneurs have among them k chains stubs of length N (i.e. $\sum_{i=1}^{l} B_N(j_i) = k$). Note that $P_N(1,k)$ is the probability an entrepreneur has exactly k supply chains of length N (i.e., the probability that $B_{j,N} = k$). We will suppress the argument j when not needed for clarity.

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 $\varphi(x) = e^{-M(1-x)}$. 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:

Lemma 21. $\varphi_N(x) = \varphi^{(N)}(x)$

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 1. *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 an entrepreneur has no supply chains that continue indefinitely.

Claim 2. *The probability that a single entrepreneur has no supply chains that continue indefinitely is the smallest root, ρ, of $y = \varphi(y)$.*

Proof. The probability that an entrepreneur has no chain stubs greater than length N is $P_N(1,0)$, or $\varphi_N(0)$. Then the probability that a single entrepreneur has no chains that continue indefinitely is $\lim_{N\to\infty} \varphi^{(N)}(0)$. Next note 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 $\left\{\varphi^{(N)}(y)\right\}$ 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 is a root of $y = \varphi(y)$, and therefore must be ρ. In other words, $\lim_{N\to\infty} \varphi^{(N)}(0) = \rho$. ■

Claim 3. *If $M \leq 1$ then with probability 1 an entrepreneur has no supply chains that continue indefinitely. If $M > 1$ then there is a strictly positive probability the entrepreneur has a supply chain that continues indefinitely.*

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

C Cross-Sectional Patterns

C.1 Distribution of Customers

Proposition 4 (1)**.** *Suppose Assumption 2 holds. Among entrepreneurs with efficiency q, the number of actual customers follows a Poisson distribution with mean $\frac{m}{\theta} q^{\alpha\zeta}$.*

Proof. We first derive an expression for $\tilde{F}(x)$. This is the probability that a potential buyer has no alternative techniques that deliver efficiency better than x. The potential buyer will have $n-1$ other techniques with probability $\frac{e^{-M}M^n}{n!(1-e^{-M})}$. The probability that a single alternative delivers efficiency no greater than x is $G(x)$. Therefore the probability that none of the potential buyer's alternatives deliver efficiency better than x is:

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

Consider an entrepreneur with efficiency q_s. If a single downstream technique has productivity z, the technique delivers efficiency zq_s^{α} to the potential customer, and will be selected by that customer with probability $\tilde{F}(zq_s^{\alpha})$. Integrating over possible productivities, the probability that a single downstream technique is used by the customer is $\int_{z_0}^{\infty} \tilde{F}(zq_s^{\alpha})dH(z)$. Since the number of downstream techniques is Poisson with mean M, the number of downstream techniques that are used is Poisson with mean $M\int_{z_0}^{\infty} \tilde{F}(zq_s^{\alpha})dH(z)$.

Using the functional form for H and taking the limit as $z_0 \to 0$ this can be simplified considerably. Since $\lim_{z_0\to 0} e^{-mz_0^{-\zeta}} = 0$ and $\lim_{z_0 \to 0} G(q) = 1$, we have that

$$\lim_{z_0\to 0} mz_0^{-\zeta} \int_{z_0}^{\infty} \frac{F(zq_s^{\alpha}) - e^{-mz_0^{-\zeta}}}{G(zq_s^{\alpha}) - e^{-mz_0^{-\zeta}}} \zeta z_0^{\zeta} z^{-\zeta-1} dz = m \int_0^{\infty} e^{-\theta(zq_s^{\alpha})^{-\zeta}} \zeta z^{-\zeta-1} dz = \frac{m}{\theta} q_s^{\alpha\zeta}$$

■

Proposition 4 (2)**.** *Suppose Assumption 2 holds. Among all entrepreneurs, the distribution of customers asymptotically follows a power law with exponent $1/\alpha$: $\sum_{k=n}^{\infty} p_k \sim \frac{1}{\Gamma(1-\alpha)^{1/\alpha}} n^{-1/\alpha}$.*

Proof. With the functional forms, among firms with efficiency q, the distribution of customers is Poisson with mean $\frac{m}{\theta}q^{\alpha\zeta}$. If $p_n(q)$ is the probability that an entrepreneur with efficiency q has n customers, $p_n(q) = \frac{\left(\frac{m}{\theta}q^{\alpha\zeta}\right)^n e^{-\frac{m}{\theta}q^{\alpha\zeta}}}{n!}$. Integrating across efficiencies, the unconditional probability that a firm has n customers is

$$p_n = \int_0^\infty p_n(q)dF(q) = \int_0^\infty \frac{\left(\frac{m}{\theta}q^{\alpha\zeta}\right)^n e^{-\frac{m}{\theta}q^{\alpha\zeta}}}{n!}dF(q)$$

We will make the change of variables $u = \frac{m}{\theta}q^{\alpha\zeta}$. Noting that $\theta = \Gamma(1-\alpha)m\theta^\alpha$, this means that $\theta q^{-\zeta} = \left(\frac{m}{\theta^{1-\alpha}}\right)^{1/\alpha}\left(\frac{m}{\theta}q^{\alpha\zeta}\right)^{-1/\alpha} = [\Gamma(1-\alpha)u]^{-1/\alpha}$ and $\zeta\frac{dq}{q} = \frac{1}{\alpha}\frac{du}{u}$. Together, these imply that $dF(q) = \zeta\theta q^{-\zeta-1}e^{-\theta q^{-\zeta}}dq = \frac{e^{-[\Gamma(1-\alpha)u]^{-1/\alpha}}}{\Gamma(1-\alpha)^{\frac{1}{\alpha}}\alpha}u^{-\frac{1}{\alpha}-1}du$, so that p_n can be written as

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

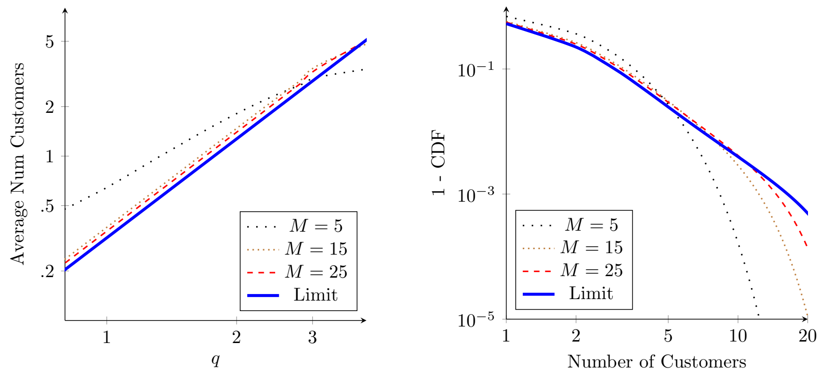
Theorem 2.1 of Willmot (1990) states that if the probabilities of a mixed Poisson distribution are given by $p_n = \int \frac{(\lambda x)^n e^{-\lambda x}}{n!}f(x)dx$, then if $f(x) \sim C(x)x^\gamma e^{-\beta x}$, $x \to \infty$ where $C(x)$ is a locally bounded function on $(0,\infty)$ which varies slowly at infinity, $\beta \geq 0$, and $-\infty < \gamma < \infty$ (with $\gamma < -1$ if $\beta = 0$) then $p_n \sim \frac{C(n)}{(\lambda+\beta)^{\gamma+1}}\left(\frac{\lambda}{\lambda+\beta}\right)^n n^\gamma$, $n \to \infty$.

Since $\lim_{u\to\infty} e^{-[\Gamma(1-\alpha)u]^{-1/\alpha}} = 1$, this theorem implies $\lim_{n\to\infty}\frac{\frac{1}{\Gamma(1-\alpha)^{1/\alpha}\alpha}n^{-\frac{1}{\alpha}-1}}{p_n} = 1$. Then Theorem 1 of Feller (1971) VIII.9 implies that $\lim_{n\to\infty}\frac{np_n}{\sum_{k=n}^\infty p_k} = \frac{1}{\alpha}$, giving the desired result. ■

Claim 4. *Under Assumption 2,* $\frac{Cov(\log q, \#\ customers)}{St.\ Dev.(\log q)} = \frac{\sqrt{6}}{\pi}\int_0^1 \frac{x^{-\alpha}-1}{1-x}dx$.

Proof. First, note that $\Gamma'(t) = \frac{d}{dt}\int_0^\infty x^{t-1}e^{-u}du = \int_0^\infty \log u u^{t-1}e^{-u}du$. Second, under Assumption 2, letting γ be the Euler-Mascheroni constant, we have

$$\begin{aligned}
\mathbb{E}[\log q] &= \int_0^\infty \log q dF(q) = \frac{1}{\zeta}\left[\log\theta - \int_0^\infty \log(\theta q^{-\zeta})dF(q)\right] = \frac{1}{\zeta}\left[\log\theta - \int_0^\infty \log u e^{-u}du\right] \\
&= \frac{1}{\zeta}[\log\theta + \gamma] \\
\mathbb{E}[(\log q)^2] &= \int_0^\infty \frac{1}{\zeta^2}\left[\log\theta - \log(\theta q^{-\zeta})\right]^2 dF(q) = \int_0^\infty \frac{1}{\zeta^2}[\log\theta - \log u]^2 e^{-u}du \\
&= \frac{1}{\zeta^2}\left[(\log\theta)^2 + 2\log\theta\gamma + \gamma^2 + \pi^2/6\right] \\
\mathbb{E}[\log q\frac{m}{\theta}q^{\alpha\zeta}] &= \int_0^\infty \frac{1}{\zeta}\left[\log\theta - \log(\theta q^{-\zeta})\right]\left(\frac{m}{\theta^{1-\alpha}}(\theta q^{-\zeta})^{-\alpha}\right)dF(q) = \int_0^\infty \frac{1}{\zeta}[\log\theta - \log u]\left(\frac{1}{\Gamma(1-\alpha)}u^{-\alpha}\right)e^{-u}du \\
&= \frac{1}{\zeta}\left[\log\theta - \frac{\int_0^\infty \log u u^{-\alpha}e^{-u}du}{\Gamma(1-\alpha)}\right] = \frac{1}{\zeta}\left[\log\theta - \frac{\Gamma'(1-\alpha)}{\Gamma(1-\alpha)}\right]
\end{aligned}$$



(a) Conditional Means (b) Unconditional Distribution

Figure 4: Distribution of Customers, Sequence of Economies

Figure 4 plots features of the distribution of customers for several economies with $H(z) = 1-(z/z_0)^{-\zeta}$ with $\zeta = 4$ and $\alpha = 0.5$. The line labeled "Limit" corresponds to an economy that satisfies Assumption 2.

The first two imply that the standard deviation of $\log q$ is $\frac{\pi}{\zeta\sqrt{6}}$. The first and third (and $\mathbb{E}[\frac{m}{\theta}q^{\alpha\zeta}] = 1$) imply that $\text{Cov}\,(\log q, \#\text{ customers}) = \frac{1}{\zeta}\left[\frac{-\Gamma'(1-\alpha)}{\Gamma(1-\alpha)} - \gamma\right] = \frac{1}{\zeta}\int_0^1 \frac{x^{-\alpha}-1}{1-x}dx$ ∎

C.1.1 Sequence of Economies

As discussed in Section 4, Assumption 2 can be interpreted as the limit of a sequence of economies in which $H(z) = 1-(z/z_0)^{-\zeta}$ and the limit as $z_0 \to 0$ is taken as $m \equiv M z_0^{-\zeta}$ is held fixed. Figure 4 shows the behavior of the distribution of customers as the sequence converges to this limit. Panel (a) shows the expected number of customers among entrepreneurs with a given level of efficiency while panel (b) plots the right CDF of the overall distribution, each on a log-log plot. As can be seen in each panel, these features of the distribution of customers converge pointwise to their respective limits, with the tail converging most slowly.

C.2 The Distribution of Employment

Because employment is the sum of several components—labor used to produce inputs for each customer and for the household—rather than working with the CDF of the size distribution, $\mathcal{L}(\cdot)$, it will be easier to work with its Laplace-Stieltjes transform, $\hat{\mathcal{L}}(s) \equiv \int_0^\infty e^{-sl}d\mathcal{L}(l)$. Similarly, if $\mathcal{L}(\cdot|q)$ is the CDF of the conditional size distribution among entrepreneurs with efficiency q, its transform is $\hat{\mathcal{L}}(s|q) \equiv \int_0^\infty e^{-sl}d\mathcal{L}(l|q)$. These are related in that $\mathcal{L}(l) = \int_0^\infty \mathcal{L}(l|q)dF(q)$ and $\hat{\mathcal{L}}(s) = \int_0^\infty \hat{\mathcal{L}}(s|q)dF(q)$. This section characterizes these transforms and then studies their implications for the size distribution.

We first derive a relationship between the conditional size distributions among entrepreneurs with different efficiencies. Recall that $\tilde{F}(q) \equiv \frac{F(q)-e^{-M}}{G(q)(1-e^{-M})}$ describes the CDF of a potential buyer's best alternative technique.

Lemma 22. *The transforms $\{\hat{\mathcal{L}}(\cdot|q)\}$ satisfy*

$$\hat{\mathcal{L}}(s|q) = e^{-s(1-\alpha)(q/Q)^{\varepsilon-1}L} e^{-M\int_0^\infty \tilde{F}(zq^\alpha)\left[1-\hat{\mathcal{L}}(\alpha s|zq^\alpha)\right]dH(z)}$$

Proof. Total labor used by an entrepreneur is the sum of labor used to make output for consumption and for use as an intermediate input by others. We use the fact that the Laplace-Stieltjes transform of a sum of random variables is the product of the transforms of each.

An entrepreneur with efficiency q uses $(1-\alpha)(q/Q)^{\varepsilon-1}L$ units of labor in making goods for the household. The transform of this is $e^{-s(1-\alpha)(q/Q)^{\varepsilon-1}L}$.

We next consider labor used to make intermediate inputs. Recall that if j uses l_j units of labor, j's supplier will use αl_j units of labor to make the inputs for j. Thus if the transform of labor used by a buyer with efficiency q_b is $\hat{\mathcal{L}}(s|q_b)$, then the transform of labor used by its supplier to make intermediates is

$$\int_0^\infty \frac{1}{\alpha}\Pr\left(l_j = \frac{l}{\alpha}\right)e^{-sl}dl = \int_0^\infty \Pr\left(l_j = \frac{l}{\alpha}\right)e^{-(\alpha s)\frac{l}{\alpha}}d\left(\frac{l}{\alpha}\right) = \hat{\mathcal{L}}(\alpha s|q_b)$$

For an entrepreneur with efficiency q, consider a single downstream technique with productivity z, so that the technique delivers efficiency to the buyer of zq^α. With probability $\tilde{F}(zq^\alpha)$ it is the buyer's best technique, in which case the transform of labor used to create intermediates for that customer is $\hat{\mathcal{L}}(\alpha s|zq^\alpha)$. With probability $1-\tilde{F}(zq^\alpha)$ the potential buyer uses an alternative supplier, in which case the transform of labor used to create intermediates for that customer is simply 1. Putting these together and integrating over possible realizations of productivity, the transform of labor used to make intermediates for a single *potential* customer is

$$\int_0^\infty \left\{\left[1-\tilde{F}(zq^\alpha)\right] + \tilde{F}(zq^\alpha)\hat{\mathcal{L}}(\alpha s|zq^\alpha)\right\}dH(z) = 1 - \int_0^\infty \tilde{F}(zq^\alpha)\left[1-\hat{\mathcal{L}}(\alpha s|zq^\alpha)\right]dH(z)$$

Each entrepreneur has n potential customers with probability $\frac{M^n e^{-M}}{n!}$, so the transform over labor used to create all intermediate goods (summing across all potential customers) is

$$\sum_{n=0}^{\infty} \frac{M^n e^{-M}}{n!} \left(1 - \int_0^{\infty} \tilde{F}(zq^{\alpha}) \left[1 - \hat{\mathcal{L}}(\alpha s|zq^{\alpha})\right] dH(z)\right)^n = e^{-M \int_0^{\infty} \tilde{F}(zq^{\alpha})\left[1-\hat{\mathcal{L}}(\alpha s|zq^{\alpha})\right]dH(z)}$$

$\hat{\mathcal{L}}(s|q)$ is simply the product of the transforms of labor used for to make final consumption and labor used to make intermediate inputs. ■

Under Assumption 2 the overall size distribution can be characterized without the intermediate step of solving for the conditional size distributions.

Lemma 23. *Define $v \equiv \frac{\varepsilon-1}{\zeta}$. Under Assumption 2, the transform $\hat{\mathcal{L}}(\cdot)$ satisfies*

$$\hat{\mathcal{L}}(s) = \int_0^{\infty} e^{-s(1-\alpha)\frac{t^{-v}}{\Gamma(1-v)}L} e^{-\frac{t^{-\alpha}}{\Gamma(1-\alpha)}\left[1-\hat{\mathcal{L}}(\alpha s)\right]} e^{-t} dt$$

Proof. First, using the functional forms, the term $M \int_0^{\infty} \tilde{F}(zq^{\alpha}) \left[1 - \hat{\mathcal{L}}(\alpha s|zq^{\alpha})\right] dH(z)$ can be written as $mz_0^{-\zeta} \int_{z_0}^{\infty} \tilde{F}(zq^{\alpha}) \left[1 - \hat{\mathcal{L}}(\alpha s|zq^{\alpha})\right] \zeta z_0^{\zeta} z^{-\zeta-1} dz$. Since $\tilde{F}(zq^{\alpha}) \to e^{\theta(zq^{\alpha})^{-\zeta}}$, this becomes (using the change of variables $w = zq^{\alpha}$):

$$\frac{mq^{\alpha\zeta}}{\theta} \int_0^{\infty} e^{\theta w^{-\zeta}} \left[1 - \hat{\mathcal{L}}(\alpha s|w)\right] \zeta\theta w^{-\zeta-1} dw = \frac{mq^{\alpha\zeta}}{\theta} \left[1 - \hat{\mathcal{L}}(\alpha s)\right]$$

where the last step follows because $e^{\theta w^{-\zeta}} \zeta\theta w^{-\zeta-1} dw = dF(w)$. We use this to express $\hat{\mathcal{L}}(\cdot)$:

$$\hat{\mathcal{L}}(s) = \int_0^{\infty} \hat{\mathcal{L}}(s|q) dF(q) \to \int_0^{\infty} e^{-s(1-\alpha)(q/Q)^{\varepsilon-1}L} e^{-\frac{mq^{\alpha\zeta}}{\theta}\left[1-\hat{\mathcal{L}}(\alpha s)\right]} \zeta\theta q^{-\zeta-1} e^{-\theta q^{-\zeta}} dq$$

The conclusion follows from the substitutions $Q^{\varepsilon-1} = \Gamma(1-v)\theta^{v}$ and $m = \frac{\theta^{1-\alpha}}{\Gamma(1-\alpha)}$, and the change of variables $t = \theta q^{-\zeta}$. ■

An immediate consequence of Lemma 23 is that the shape of the size distribution depends only on α and $\frac{\varepsilon}{\zeta}$.

C.3 Tail Behavior

Let $\rho = \min\left\{\alpha^{-1}, v^{-1}\right\}$ and let N be the greatest integer that is strictly less then ρ. For any integer n, let $\mu_n \equiv \int_0^{\infty} l^n d\mathcal{L}(l) = (-1)^n \hat{\mathcal{L}}^{(n)}(0)$ be the nth moment of the size distribution. The strategy is to show that $\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(s)$ is regularly varying with index $\rho - N$ as $s \searrow 0$. Using the Tauberian theorem of Bingham and Doney (1974), this will imply that $1 - \mathcal{L}(l)$ is regularly varying with index $-\rho$ as $l \to \infty$. The theorem gives this implication only when $0 < \rho - N < 1$, so we restrict attention to that case.

Define $\varphi(s;t) \equiv s\frac{(1-\alpha)}{\Gamma(1-v)}t^{-v} + \frac{1-\hat{\mathcal{L}}(\alpha s)}{\Gamma(1-\alpha)}t^{-\alpha}$ so that $\hat{\mathcal{L}}(s) = \int_0^\infty e^{-\varphi(s;t)}e^{-t}dt$. Since we will be interested in $\hat{\mathcal{L}}^{(n)}$, it will be useful to derive an expression for $\frac{d^n}{ds^n}\left[e^{-\varphi(s;t)}\right]$. By Faa di Bruno's formula (a generalization of the chain rule to higher derivatives), we have that

$$\frac{d^n}{ds^n}\left[e^{-\varphi(s;t)}\right] = e^{-\varphi(s;t)} \sum_{\iota \in I_n} \frac{n!}{\iota_1!(1!)^{\iota_1}...\iota_n!(n!)^{\iota_n}} \prod_{j=1}^{n}\left[-\varphi^{(j)}(s;t)\right]^{\iota_j}$$

where I_n is the set of all n-tuples of non-negative integers $\iota = (\iota_1, ..., \iota_n)$ such that $1\iota_1 + ... + n\iota_n = n$.

Since $\varphi^{(1)}(s;t) = \frac{(1-\alpha)}{\Gamma(1-v)}t^{-v} + \frac{-\alpha\hat{\mathcal{L}}^{(1)}(\alpha s)}{\Gamma(1-\alpha)}t^{-\alpha}$ and $\varphi^{(j)}(s;t) = \frac{-\alpha^j\hat{\mathcal{L}}^{(j)}(\alpha s)}{\Gamma(1-\alpha)}t^{-\alpha}$ for $j \geq 2$, we have for each $\iota \in I_n$ that

$$\begin{aligned}
\prod_{j=1}^{n}\left[-\varphi^{(j)}(s;t)\right]^{\iota_j} &= \sum_{k=0}^{\iota_1}\binom{\iota_1}{k}\left[\frac{-(1-\alpha)}{\Gamma(1-v)}t^{-v}\right]^k\left[\frac{\alpha\hat{\mathcal{L}}^{(1)}(\alpha s)}{\Gamma(1-\alpha)}t^{-\alpha}\right]^{\iota_1-k}\prod_{j=2}^{n}\left[\frac{\alpha^j\hat{\mathcal{L}}^{(j)}(\alpha s)}{\Gamma(1-\alpha)}t^{-\alpha}\right]^{\iota_j} \\
&= \sum_{k=0}^{\iota_1}\binom{\iota_1}{k}\left[\frac{-(1-\alpha)}{\Gamma(1-v)}\frac{\Gamma(1-\alpha)}{\alpha\hat{\mathcal{L}}^{(1)}(\alpha s)}\right]^k\prod_{j=1}^{n}\left[\frac{\alpha^j\hat{\mathcal{L}}^{(j)}(\alpha s)}{\Gamma(1-\alpha)}\right]^{\iota_j} t^{-\left[\alpha\left(\sum_{j=1}^n \iota_j - k\right)+vk\right]}
\end{aligned}$$

Thus we can write

$$\frac{d^n}{ds^n}\left[e^{-\varphi(s;t)}\right] = e^{-\varphi(s;t)}\sum_{\iota\in I_n}\sum_{k=0}^{\iota_1} B_{n,\iota,k}(s)t^{-\beta(\iota,k)} \tag{34}$$

where

$$\begin{aligned}
B_{n,\iota,k}(s) &\equiv \frac{n!}{\iota_1!(1!)^{\iota_1}...\iota_n!(n!)^{\iota_n}}\binom{\iota_1}{k}\left[\frac{-(1-\alpha)}{\Gamma(1-v)}\frac{\Gamma(1-\alpha)}{\alpha\hat{\mathcal{L}}^{(1)}(\alpha s)}\right]^k\prod_{j=1}^{n}\left[\frac{\alpha^j\hat{\mathcal{L}}^{(j)}(\alpha s)}{\Gamma(1-\alpha)}\right]^{\iota_j} \\
\beta(\iota,k) &\equiv \alpha\left(\sum_{j=1}^{n}\iota_j - k\right) + vk
\end{aligned}$$

Note that if $n < \rho$ then each $\beta(\iota, k) \in (0,1)$. We first show that the nth derivative can be taken inside the integral

Lemma 24. *For any integer $n < \rho$, $\hat{\mathcal{L}}^{(n)}(s) = \int_0^\infty \frac{d^n}{ds^n}\left[e^{-\varphi(s;t)}\right]e^{-t}dt$.*

Proof. We first show that for each $\iota \in I_{n-1}$,

$$\frac{d}{ds}\int_0^\infty e^{-\varphi(s;t)}t^{-\beta(\iota,k)}e^{-t}dt = \int_0^\infty \frac{d}{ds}\left[e^{-\varphi(s;t)}\right]t^{-\beta(\iota,k)}e^{-t}dt$$

Since $\frac{d}{ds}\left[e^{-\varphi(s;t)}\right] = \left(\frac{(1-\alpha)}{\Gamma(1-v)}t^{-v} + \frac{-\alpha\hat{\mathcal{L}}^{(1)}(\alpha s)}{\Gamma(1-\alpha)}t^{-\alpha}\right)e^{-\varphi(s;t)}$, the integrand on RHS is dominated by $\left(\frac{(1-\alpha)}{\Gamma(1-v)}t^{-v} + \frac{-\alpha\hat{\mathcal{L}}^{(1)}(\alpha s)}{\Gamma(1-\alpha)}t^{-\alpha}\right)t^{-\beta(\iota,k)}e^{-t}$, which is integrable for each s because $\alpha + \beta(\iota,k) < 1$ and $v + \beta(\iota,k) < 1$.

With this, we proceed by induction. Trivially, the conclusion holds for $n = 0$. Now for $n < \rho$, assume that $\hat{\mathcal{L}}^{(n-1)}(s) = \int_0^\infty \frac{d^{n-1}}{ds^{n-1}}\left[e^{-\varphi(s;t)}\right]e^{-t}dt$. Equation (34) implies

$$\hat{\mathcal{L}}^{(n-1)}(s) = \sum_{\iota\in I_{n-1}}\sum_{k=0}^{\iota_1} B_{n-1,\iota,k}(s)\int_0^\infty e^{-\varphi(s;t)}t^{-\beta(\iota,k)}e^{-t}dt$$

Differentiating each side gives

$$\begin{aligned}\hat{\mathcal{L}}^{(n)}(s) &= \sum_{\iota\in I_{n-1}}\sum_{k=0}^{\iota_1}\frac{dB_{n-1,\iota,k}(s)}{ds}\int_0^\infty e^{-\varphi(s;t)}t^{-\beta(\iota,k)}e^{-t}dt + B_{n-1,\iota,k}(s)\frac{d}{ds}\int_0^\infty e^{-\varphi(s;t)}t^{-\beta(\iota,k)}e^{-t}dt\\ &= \int_0^\infty \sum_{\iota\in I_{n-1}}\sum_{k=0}^{\iota_1}\left[\frac{dB_{n-1,\iota,k}(s)}{ds}e^{-\varphi(s;t)} + B_{n-1,\iota,k}(s)\frac{de^{-\varphi(s;t)}}{ds}\right]t^{-\beta(\iota,k)}e^{-t}dt\\ &= \int_0^\infty \frac{d^n}{ds^n}\left[e^{-\varphi(s;t)}\right]e^{-t}dt\end{aligned}$$

■

Lemma 25. *For any integer $n < \rho$, $\mu_n < \infty$.*

Proof. Again, we proceed by induction. $\mu_0 = 1$. Now assume that $\mu_0, ..., \mu_{n-1} < \infty$. We begin with the expression for $\hat{\mathcal{L}}^{(n)}(s)$:

$$\hat{\mathcal{L}}^{(n)}(s) = \sum_{\iota\in I_n}\sum_{k=0}^{\iota_1} B_{n,\iota,k}(s)\int_0^\infty e^{-\varphi(s;t)}t^{-\beta(\iota,k)}e^{-t}dt \tag{35}$$

Define $\tilde{I}_n \equiv I_n \setminus (0, ..., 0, 1)$. Then pulling out from the sum the term for $\iota = (0, ..., 0, 1)$, we have

$$\hat{\mathcal{L}}^{(n)}(s) = \frac{\alpha^n\hat{\mathcal{L}}^{(n)}(\alpha s)}{\Gamma(1-\alpha)}\int_0^\infty e^{-\varphi(s;t)}t^{-\alpha}e^{-t}dt + \sum_{\iota\in\tilde{I}_n}\sum_{k=0}^{\iota_1} B_{n,\iota,k}(s)\int_0^\infty e^{-\varphi(s;t)}t^{-\beta(\iota,k)}e^{-t}dt$$

We can take the limit as $s \searrow 0$ of each side. Since $e^{-\varphi(s;t)}$ is dominated by 1, the limit can be taken inside of each integral, and since $\lim_{s\searrow 0}\varphi(s;t) = 0$, we have

$$\hat{\mathcal{L}}^{(n)}(0) = \alpha^n\hat{\mathcal{L}}^{(n)}(0) + \sum_{\iota\in\tilde{I}_n}\sum_{k=0}^{\iota_1} B_{n,\iota,k}(0)\Gamma\left\{1-\beta(\iota,k)\right\}$$

For each $\iota \in \tilde{I}_n$, and for each k, $B_{n,\iota,k}(0)$ is proportional to a product of derivatives of $\hat{\mathcal{L}}$, and each of those derivatives is of order less than n. Since all of these are finite, $\hat{\mathcal{L}}^{(n)}(0) < \infty$. ■

With this we show that $\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(s)$ is regularly varying as $s \searrow 0$. First, we have

$$\begin{aligned}
\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(s) &= (-1)^N \left[\hat{\mathcal{L}}^{(N)}(0) - \hat{\mathcal{L}}^{(N)}(s)\right] \\
&= (-1)^N \sum_{\iota \in I_n} \sum_{k=0}^{\iota_1} B_{n,\iota,k}(0) \int_0^\infty t^{-\beta(\iota,k)} e^{-t} dt - B_{n,\iota,k}(s) \int_0^\infty e^{-\varphi(s;t)} t^{-\beta(\iota,k)} e^{-t} dt
\end{aligned}$$

We can decompose the object of interest into three terms:

$$\frac{\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(s)}{s^{\rho - N}} = A_1(s) + A_2(s) + A_3(s)$$

where A_1, A_2 and A_3 are defined as

$$\begin{aligned}
A_1(s) &\equiv s^{N-\rho}(-1)^N \sum_{\iota \in I_N} \sum_{k=0}^{\iota_1} \left[B_{N,\iota,k}(0) - B_{N,\iota,k}(s)\right] \int_0^\infty t^{-\beta(\iota,k)} e^{-t} dt \\
A_2(s) &\equiv s^{N-\rho}(-1)^N \sum_{\iota \in I_N} \sum_{k=0}^{\iota_1} B_{N,\iota,k}(s) \int_0^\infty t^{-\beta(\iota,k)} \left[1 - e^{-\varphi(s;t)}\right] dt \\
A_3(s) &\equiv s^{N-\rho}(-1)^{N+1} \sum_{\iota \in I_N} \sum_{k=0}^{\iota_1} B_{N,\iota,k}(s) \int_0^\infty t^{-\beta(\iota,k)} \left[1 - e^{-\varphi(s;t)}\right] \left[1 - e^{-t}\right] dt
\end{aligned}$$

This particular decomposition is useful because it will allow for the use of the Monotone Convergence Theorem in characterizing the limiting behavior of A_2 and A_3.

Lemma 26. *If $\rho \notin \mathbb{N}$, $\lim_{s \searrow 0} A_1(s) = \alpha^\rho \lim_{s \to 0} \frac{\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(s)}{s^{\rho - N}}$.*

Proof. As above, we can separate the term for $\iota = (0, ..., 0, 1)$ from $\tilde{I}_N$, to write

$$\begin{aligned}
\lim_{s \searrow 0} A_1(s) &= \lim_{s \searrow 0} s^{N-\rho}(-1)^N \sum_{\iota \in I_N} \sum_{k=0}^{\iota_1} \left[B_{N,\iota,k}(0) - B_{N,\iota,k}(s)\right] \Gamma\left(1 - \beta(\iota,k)\right) \\
&= \lim_{s \searrow 0} (-1)^N \frac{\left[\frac{\alpha^N \hat{\mathcal{L}}^{(N)}(0)}{\Gamma(1-\alpha)} - \frac{\alpha^N \hat{\mathcal{L}}^{(N)}(\alpha s)}{\Gamma(1-\alpha)}\right]}{s^{\rho - N}} \Gamma(1-\alpha) \\
&\quad + \lim_{s \searrow 0} (-1)^N \sum_{\iota \in \tilde{I}_N} \sum_{k=0}^{\iota_1} \frac{\left[B_{N,\iota,k}(0) - B_{N,\iota,k}(s)\right]}{s^{\rho - N}} \Gamma\left(1 - \beta(\iota,k)\right)
\end{aligned}$$

The first term is simply $\alpha^\rho \lim_{s \searrow 0} \frac{\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(\alpha s)}{(\alpha s)^{\rho - N}}$. The second term equals zero: After using L'Hospital's rule, the numerator becomes a mulitnomial of derivatives of $\hat{\mathcal{L}}$ or order no greater than

N, (so all of these are finite) while the denominator becomes $(\rho-N)s^{\rho-N-1}$ which goes to infinity. ■

To characterize the limiting behavior of A_2 and A_3, it will be useful to define

$$\begin{aligned}\kappa &\equiv \frac{1}{\Gamma(1-\rho^{-1})}\left((1-\alpha)\mathbb{I}_{v\geq\alpha}+\alpha\mathbb{I}_{\alpha\geq v}\right)\\ \tilde{\varphi}(s;w) &\equiv \varphi\left(s;\left[1-\hat{\mathcal{L}}(\alpha s)\right]^{\rho}\kappa^{\rho}\alpha^{-\rho}w^{-\rho}\right)\end{aligned}$$

$\tilde{\varphi}$ is defined this way so that $\lim_{s\to 0}\tilde{\varphi}(s;w)=w$ (this can be easily verified for each of the three cases: $\alpha>v$, $\alpha=v$, and $\alpha<v$).

Lemma 27. *$\tilde{\varphi}(s;w)$ is non-decreasing in s in the neighborhood of zero.*

Proof. Using the definitions of $\tilde{\varphi}$ and φ, we have

$$\tilde{\varphi}(s;w)=s^{1-\rho v}(1-\alpha)\frac{\kappa^{-\rho v}\left[\frac{1-\hat{\mathcal{L}}(\alpha s)}{\alpha s}\right]^{-\rho v}}{\Gamma(1-v)}w^{\rho v}+\frac{(\kappa/\alpha)^{-\rho\alpha}}{\Gamma(1-\alpha)}\left[1-\hat{\mathcal{L}}(\alpha s)\right]^{1-\alpha\rho}w^{\alpha\rho}$$

We first show that $\frac{1-\hat{\mathcal{L}}(s)}{s}$ is non-increasing in a neighborhood of zero. Since $\mu_2>0$,

$$\lim_{s\searrow 0}\frac{d}{ds}\left(\frac{1-\hat{\mathcal{L}}(s)}{s}\right)=\lim_{s\searrow 0}\frac{-s\hat{\mathcal{L}}^{(1)}(s)-\left[1-\hat{\mathcal{L}}(s)\right]}{s^2}=-\hat{\mathcal{L}}^{(2)}(0)<0 \tag{36}$$

Next, since $1-\hat{\mathcal{L}}(\alpha s)$ is non-decreasing in s, and since both $1-\rho v$ and $1-\rho\alpha$ are non-negative, $\tilde{\varphi}(s;w)$ is nondecreasing in the neighborhood of zero. ■

Lemma 28. $\lim_{s\searrow 0}A_2(s)=\kappa^{\rho}\frac{\rho}{\rho-N}\Gamma(1-\rho+N)$.

Proof. Using the change of variables $w=\left[1-\hat{\mathcal{L}}(\alpha s)\right]\kappa\alpha^{-1}t^{-1/\rho}$, we have

$$A_2(s)=(-1)^N\sum_{\iota\in I_N}\sum_{k=0}^{\iota_1}s^{N-\rho\beta(\iota,k)}B_{N,\iota,k}(s)\left\{\frac{\left[1-\hat{\mathcal{L}}(\alpha s)\right]}{\alpha s}\kappa\right\}^{\rho-\rho\beta(\iota,k)}\int_0^{\infty}\left[1-e^{-\tilde{\varphi}(s;w)}\right]\rho w^{-\rho[1-\beta(\iota,k)]-1}dw$$

We next take the limit of each side as s goes to zero. Lemma 27 and the monotone convergence theorem imply that the limit can be brought inside the integral to yield

$$\lim_{s\searrow 0}A_2(s)=(-1)^N\sum_{\iota\in I_N}\sum_{k=0}^{\iota_1}\left(\lim_{s\searrow 0}s^{N-\rho\beta(\iota,k)}\right)B_{N,\iota,k}(0)\kappa^{\rho-\rho\beta(\iota,k)}\int_0^{\infty}\left[1-e^{-w}\right]\rho w^{-\rho[1-\beta(\iota,k)]-1}dw$$

Noting that $N \geq \rho\beta(\iota, k)$, the term $\lim_{s\to 0} s^{N-\rho\beta(\iota,k)}$ is zero unless $N = \rho\beta(\iota, k)$. Thus $\lim_{s\searrow 0} A_2(s)$ can be written as

$$\lim_{s\searrow 0} A_2(s) = (-1)^N \kappa^{\rho-N} \int_0^\infty \left[1 - e^{-w}\right] \rho w^{-(\rho-N)-1} dw \sum_{\iota\in I_N} \sum_{k=0}^{\iota_1} B_{N,\iota,k}(0)\mathbb{I}_{N=\rho\beta(\iota,k)}$$

The integral is $\int_0^\infty \left[1 - e^{-w}\right] \rho w^{-(\rho-N)-1} dw = \frac{\rho}{\rho-N}\Gamma(1 - \rho + N)$. To finish the proof, we show that

$$(-1)^N \sum_{\iota\in I_N} \sum_{k=0}^{\iota_1} B_{N,\iota,k}(0)\mathbb{I}_{N=\rho\beta(\iota,k)} = \kappa^N \tag{37}$$

To see this note first that $N = \rho\beta(\iota, k)$ requires $\iota = (N, 0, ..., 0)$. If $\alpha > \upsilon$, $N = \rho\beta(\iota, k)$ also requires $k = 0$, whereas if $\alpha < \upsilon$, it requires $k = N$. If $\alpha = \upsilon$, $N = \rho\beta(\iota, k)$ for each $k \in \{0, ..., N\}$. For each of these three cases, one can compute each non-zero term in the sum and verify equation (37).

■

Lemma 29. $\lim_{s\searrow 0} A_3(s) = 0$

Proof. The strategy is the same as in the previous lemma. Using the same change of variables $w = \left[1 - \hat{\mathcal{L}}(\alpha s)\right] \kappa\alpha^{-1} t^{-1/\rho}$, we have

$$\begin{aligned} A_3(s) \quad &= \quad (-1)^N \sum_{\iota\in I_N} \sum_{k=0}^{\iota_1} s^{N-\rho\beta(\iota,k)} B_{N,\iota,k}(s) \left\{ \frac{\left[1 - \hat{\mathcal{L}}(\alpha s)\right]}{\alpha s} \kappa \right\}^{\rho-\rho\beta(\iota,k)} \\ &\quad \times \int_0^\infty \left[1 - e^{-\tilde{\varphi}(s;w)}\right] \left[1 - e^{-\left[1-\hat{\mathcal{L}}(\alpha s)\right]^\rho (\kappa/\alpha)^\rho w^{-\rho}}\right] \rho w^{-\rho[1-\beta(\iota,k)]-1} dw \end{aligned}$$

We can take a limit of each side. Since both $\tilde{\varphi}(s; w)$ and $\left[1 - \hat{\mathcal{L}}(\alpha s)\right]^\rho (\kappa/\alpha)^\rho w^{-\rho}$ are non-decreasing in s in the neighborhood of $s = 0$, we can use the monotone convergence theorem to bring the limit inside the integral. Thus we have

$$\begin{aligned} \lim_{s\searrow 0} A_3(s) \quad &= \quad (-1)^N \sum_{\iota\in I_N} \sum_{k=0}^{\iota_1} \left(\lim_{s\searrow 0} s^{N-\rho\beta(\iota,k)}\right) B_{N,\iota,k}(0)\kappa^{\rho-\rho\beta(\iota,k)} \\ &\quad \times \int_0^\infty \lim_{s\searrow 0} \left[1 - e^{-\tilde{\varphi}(s;w)}\right] \left[1 - e^{-\left[1-\hat{\mathcal{L}}(\alpha s)\right]^\rho (\kappa/\alpha)^\rho w^{-\rho}}\right] \rho w^{-\rho[1-\beta(\iota,k)]-1} dw \end{aligned}$$

Since the limit of the integrand is zero for each integral, we have $\lim_{s\searrow 0} A_3(s) = 0$. ■

We finally come to the main result.

Proposition 5 (1)**.** *Suppose that Assumption 2 holds and that* $\rho \equiv \min\left\{\frac{1}{\alpha}, \frac{1}{(\varepsilon-1)/\zeta}\right\}$ *is not an integer. Then* $1-\mathcal{L}(l) \sim \frac{\kappa^\rho}{1-\alpha^\rho} l^{-\rho}$

Proof. The previous lemmas imply that

$$\lim_{s\searrow 0} \frac{\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(s)}{s^{\rho-N}} = \alpha^\rho \lim_{s\searrow 0} \frac{\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(s)}{s^{\rho-N}} + \kappa^\rho \frac{\rho}{\rho-N}\Gamma(1-\rho+N) + 0$$

or

$$\lim_{s\searrow 0} \frac{\mu_N - (-1)^N \hat{\mathcal{L}}^{(N)}(s)}{s^{\rho-N}} = \frac{\kappa^\rho}{1-\alpha^\rho}\frac{\rho}{\rho-N}\Gamma(1-\rho+N)$$

By Theorem A in Bingham and Doney (1974) we therefore have that

$$\lim_{l\to\infty} \frac{1-\mathcal{L}(l)}{l^{-\rho}} = \frac{\kappa^\rho}{1-\alpha^\rho}\left[(-1)^N \frac{\Gamma(\rho-N)}{\Gamma(\rho)}\frac{\Gamma(1-\rho+N)}{\Gamma(1-\rho)}\right]$$

Since $\Gamma(\rho) = \Gamma(\rho-N)\prod_{k=1}^N(\rho-k)$ and $(-1)^N\Gamma(1-\rho+N) = \Gamma(1-\rho)\prod_{k=1}^N(-1)(1-\rho+N-k) = \Gamma(1-\rho)\prod_{k=1}^N(\rho-k)$, the term in brackets equals unity, completing the proof. ■

Next we turn to the tail behavior of the conditional size distribution, $\mathcal{L}(\cdot|q)$.

Proposition 5 (2)**.** *Suppose that Assumption 2 holds and that* $\rho \equiv \min\left\{\frac{1}{\alpha}, \frac{1}{(\varepsilon-1)/\zeta}\right\}$ *is not an integer. Then* $1-\mathcal{L}(l|q) \sim \frac{mq^{\alpha\zeta}}{\theta}\alpha^\rho\left[1-\mathcal{L}(l)\right]$

Proof. Following the logic of Lemma 23, the transform of $\mathcal{L}(\cdot|q)$ can be written as $\hat{\mathcal{L}}(s|q) = e^{-\varphi\left(s;\theta q^{-\zeta}\right)}$, with derivatives

$$\hat{\mathcal{L}}^{(n)}(s|q) = e^{-\varphi\left(s,\theta q^{-\zeta}\right)}\sum_{\iota\in I_n}\sum_{k=0}^{\iota_1} B_{n,\iota,k}(s)\left(\theta q^{-\zeta}\right)^{-\beta(\iota,k)}$$

For $k \le n < \rho$, $B_{n,\iota,k}(0)$ is finite, so $\mu_n(q) = \hat{\mathcal{L}}^{(n)}(0|q) < \infty$ for each $n < \rho$. Then, using $t(q) \equiv \theta q^{-\zeta}$, we have

$$\begin{aligned}
\lim_{s\searrow 0}\frac{\hat{\mathcal{L}}^{(N)}(0|q) - \hat{\mathcal{L}}^{(N)}(s|q)}{s^{\rho-N}} &= \lim_{s\searrow 0}\frac{\sum_{\iota\in I_n}\sum_{k=0}^{\iota_1} B_{n,\iota,k}(0)t(q)^{-\beta(\iota,k)} - e^{-\varphi(s,t(q))}\sum_{\iota\in I_n}\sum_{k=0}^{\iota_1} B_{n,\iota,k}(s)t(q)^{-\beta(\iota,k)}}{s^{\rho-N}} \\
&= \lim_{s\searrow 0}\sum_{\iota\in I_n}\sum_{k=0}^{\iota_1}\frac{B_{n,\iota,k}(0)-B_{n,\iota,k}(s)}{s^{\rho-N}}t(q)^{-\beta(\iota,k)} \\
&+ \lim_{s\searrow 0}\frac{1-e^{-\varphi(s,t(q))}}{s^{\rho-N}}\sum_{\iota\in I_n}\sum_{k=0}^{\iota_1} B_{n,\iota,k}(s)t(q)^{-\beta(\iota,k)}
\end{aligned}$$

Using the logic of Lemma 26, the first term is

$$\begin{aligned}
\lim_{s\searrow 0}\sum_{\iota\in I_n}\sum_{k=0}^{\iota_1}\frac{B_{n,\iota,k}(0)-B_{n,\iota,k}(s)}{s^{\rho-N}}t^{-\beta(\iota,k)} &= \lim_{s\searrow 0}\frac{\frac{\alpha^N\hat{\mathcal{L}}^{(N)}(0)}{\Gamma(1-\alpha)}-\frac{\alpha^N\hat{\mathcal{L}}^{(N)}(\alpha s)}{\Gamma(1-\alpha)}}{s^{\rho-N}}t^{-\alpha} \\
&\quad +\lim_{s\searrow 0}\sum_{\iota\in\tilde{I}_n}\sum_{k=0}^{\iota_1}\frac{\left[B_{n,\iota,k}(0)t^{-\beta(\iota,k)}-B_{n,\iota,k}(s)\right]}{s^{\rho-N}}t^{-\beta(\iota,k)} \\
&= \frac{\alpha^\rho t^{-\alpha}}{\Gamma(1-\alpha)}\lim_{s\searrow 0}\frac{\hat{\mathcal{L}}^{(N)}(0)-\hat{\mathcal{L}}^{(N)}(s)}{s^{\rho-N}}
\end{aligned}$$

The second term is zero because, using L'Hospital gives

$$\lim_{s\searrow 0}\frac{1-e^{-\varphi(s,t)}}{s^{\rho-N}}=\lim_{s\searrow 0}\frac{\varphi^{(1)}\left(s,t\right)e^{-\varphi(s,t)}}{\left(\rho-N\right)s^{\rho-N-1}}=0$$

We thus have

$$\lim_{s\searrow 0}\frac{\hat{\mathcal{L}}^{(N)}\left(0|q\right)-\hat{\mathcal{L}}^{(N)}\left(s|q\right)}{s^{\rho-N}}=\frac{\alpha^\rho\left(\theta q^{-\zeta}\right)^{-\alpha}}{\Gamma(1-\alpha)}\lim_{s\searrow 0}\frac{\hat{\mathcal{L}}^{(N)}(0)-\hat{\mathcal{L}}^{(N)}(s)}{s^{\rho-N}}$$

The result follows from Theorem A in Bingham and Doney (1974) and noting that $\frac{m}{\theta}=\frac{\theta^{-\alpha}}{\Gamma(1-\alpha)}$. ■

C.4 The Cost Share of Intermediate Inputs

Here we fill in the missing step from the proof of Proposition 7. We show that if

$$\rho\equiv\theta\int_0^\infty\int_0^\infty\left[V(\max\{u,\tilde{q}\})-V(\tilde{q})\right]dF(\tilde{q})\zeta u^{-\zeta-1}du$$

and $V(q)=\frac{1}{\varepsilon-1}(q/Q)^{\varepsilon-1}wL+\frac{m}{\theta}q^{\alpha\zeta}\rho$ imply that $\rho=\frac{1}{1-\alpha}\frac{1}{\zeta}wL$.

To see this, starting with the expression for ρ, we can integrate the inner integral by parts and switch the order of integration to get

$$
\begin{aligned}
\rho &= \theta \int_0^\infty \int_0^\infty \left[V\left(\max\{u, \tilde{q}\}\right) - V\left(\tilde{q}\right)\right] d\tilde{F}\left(\tilde{q}\right) \zeta u^{-\zeta-1} du \\
&= \theta \int_0^\infty \int_0^u \left[V\left(u\right) - V\left(\tilde{q}\right)\right] d\tilde{F}\left(\tilde{q}\right) \zeta u^{-\zeta-1} du \\
&= \theta \int_0^\infty \left[\left[V\left(u\right) - V\left(\tilde{q}\right)\right] \tilde{F}\left(\tilde{q}\right)\Big|_0^u - \int_0^u -V'\left(\tilde{q}\right) \tilde{F}\left(\tilde{q}\right) d\tilde{q}\right] \zeta u^{-\zeta-1} du \\
&= \theta \int_0^\infty \int_0^u V'\left(\tilde{q}\right) \tilde{F}\left(\tilde{q}\right) d\tilde{q} \zeta u^{-\zeta-1} du \\
&= \theta \int_0^\infty \int_{\tilde{q}}^\infty \zeta u^{-\zeta-1} du V'\left(\tilde{q}\right) \tilde{F}\left(\tilde{q}\right) d\tilde{q} \\
&= \theta \int_0^\infty \tilde{q}^{-\zeta} V'\left(\tilde{q}\right) \tilde{F}\left(\tilde{q}\right) d\tilde{q}
\end{aligned}
$$

Using the expressions for V and $\tilde{F}$, we have

$$
\begin{aligned}
\rho &= \theta \int_0^\infty \tilde{q}^{-\zeta} V'\left(\tilde{q}\right) \tilde{F}\left(\tilde{q}\right) d\tilde{q} \\
&= \theta \int_0^\infty \tilde{q}^{-\zeta} \left[\frac{q^{\varepsilon-2}}{Q^{\varepsilon-1}} wL + \frac{m}{\theta} \alpha\zeta \tilde{q}^{\alpha\zeta-1} \rho\right] e^{-\theta \tilde{q}^{-\zeta}} d\tilde{q} \\
&= \frac{1}{\zeta} \int_0^\infty \left[\frac{\theta^{\frac{\varepsilon-1}{\zeta}} x^{-\frac{\varepsilon-1}{\zeta}}}{Q^{\varepsilon-1}} wL + \frac{m}{\theta^{1-\alpha}} \alpha\zeta x^{-\alpha} \rho\right] e^{-x} dx \\
&= \frac{1}{\zeta} \left[\frac{\theta^{\frac{\varepsilon-1}{\zeta}} \Gamma\left(1 - \frac{\varepsilon-1}{\zeta}\right)}{Q^{\varepsilon-1}} wL + \frac{m}{\theta^{1-\alpha}} \Gamma\left(1-\alpha\right) \alpha\zeta\rho\right] \\
&= \frac{1}{\zeta} \left[wL + \alpha\zeta\rho\right]
\end{aligned}
$$

Solving gives for ρ gives $\rho = \frac{1}{\zeta(1-\alpha)} wL$

C.4.1 Value Added

We next provide and expression for average value added among those with efficiency q. We first compute the surplus of the techniques used by those with efficiency q, $\int_{\{j|q_j=q\}} \sum_{\phi \in U_j} \mathcal{S}\left(\phi\right) dj$. To do this, define $\tilde{F}\left(\tilde{q}|q\right)$ be the CDF of the distribution among those with efficiency $\tilde{q}$ of the efficiency delivered by the next best option. The average surplus from upstream techniques is thus

$$
\int_{\{j|q_j=q\}} \sum_{\phi \in U_j} \mathcal{S}\left(\phi\right) dj = \int_0^q \left[V\left(q\right) - V\left(\tilde{q}\right)\right] d\tilde{F}\left(\tilde{q}|q\right)
$$

Integrating by parts yields

$$\begin{aligned}\int_{\{j|q_j=q\}}\sum_{\phi\in U_j}\mathcal{S}(\phi)dj &= \left[V(q)-V(\tilde{q})\right]\tilde{F}(\tilde{q}|q)\Big|_0^q+\int_0^q V'(\tilde{q})\tilde{F}(\tilde{q}|q)\\ &= \int_0^q V'(\tilde{q})\tilde{F}(\tilde{q}|q)\end{aligned}$$

To find $\tilde{F}(\tilde{q}|q)$, we first find the joint distribution of q and $\tilde{q}$. For $\tilde{q}<q$, the probability that $\tilde{q}_j<\tilde{q}$ and $q_j<q$ if j has n techniques is $G(\tilde{q})^n+n\left[G(q)-G(\tilde{q})\right]G(\tilde{q})^{n-1}$. This the sum of the probabilities that both q_j and $\tilde{q}_j$ are no greater than $\tilde{q}$ and the probability that one of the n techniques delivers efficiency between $\tilde{q}$ and q and the other n deliver efficiency weakly less than $\tilde{q}$. Summing over techniques, we have the probability that $\tilde{q}_j<\tilde{q}$ and $q_j<q$ is

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{e^{-M}M^n}{n!}\left\{G(\tilde{q})^n+n\left[G(q)-G(\tilde{q})\right]G(\tilde{q})^{n-1}\right\} &= F(\tilde{q})+M\left[G(q)-G(\tilde{q})\right]F(\tilde{q})\\ &\rightarrow \left[1+\left(\theta\tilde{q}^{-\zeta}-\theta q^{-\zeta}\right)\right]e^{-\theta\tilde{q}^{-\zeta}}\end{aligned}$$

Thus using Bayes rule, $\tilde{F}(\tilde{q}|q)=\frac{\Pr(\tilde{q}_j\leq\tilde{q},q_j=q)}{\Pr(q_j=q)}=\tilde{F}(\tilde{q}|q)=\frac{e^{-\theta\tilde{q}^{-\zeta}}}{e^{-\theta q^{-\zeta}}}$. Using this along with the expression for $V(q)$ gives

$$\int_{\{j|q_j=q\}}\sum_{\phi\in U_j}\mathcal{S}(\phi)\,dj=\int_0^q\left[(\tilde{q}/Q)^{\varepsilon-1}+\frac{\alpha}{1-\alpha}\frac{m}{\theta}\tilde{q}^{\alpha\zeta}\right]e^{-\theta\left[\tilde{q}^{-\zeta}-q^{-\zeta}\right]}\frac{d\tilde{q}}{\tilde{q}}wL$$

Value added for entrepreneur j is

$$\pi_j^0+\beta r_j+\frac{wl_j}{1-\alpha}-\frac{\alpha}{1-\alpha}wl_j-\beta\sum_{\phi\in U_j}\mathcal{S}(\phi)\,dj$$

so that integrating over those with efficiency q gives

$$
\begin{aligned}
& \frac{1}{\varepsilon-1}\left(q/Q\right)^{\varepsilon-1}wL+\beta R\left(q\right)+w\left\{\left(1-\alpha\right)\left(q/Q\right)^{\varepsilon-1}+\alpha\frac{m}{\theta}q^{\alpha\zeta}\right\}L \\
& -\beta\int_0^q\left[\left(\tilde{q}/Q\right)^{\varepsilon-1}+\frac{\alpha}{1-\alpha}\frac{m}{\theta}\tilde{q}^{\alpha\zeta}\right]e^{-\theta\left[\tilde{q}^{-\zeta}-q^{-\zeta}\right]}\frac{d\tilde{q}}{\tilde{q}}wL \\
= \; & \frac{1}{\varepsilon-1}\left(q/Q\right)^{\varepsilon-1}wL+\beta\frac{m}{\theta}q^{\alpha\zeta}\frac{1}{1-\alpha}\frac{1}{\zeta}wL+w\left\{\left(1-\alpha\right)\left(q/Q\right)^{\varepsilon-1}+\alpha\frac{m}{\theta}q^{\alpha\zeta}\right\}L \\
& -\beta\int_0^q\left[\left(\tilde{q}/Q\right)^{\varepsilon-1}+\frac{\alpha}{1-\alpha}\frac{m}{\theta}\tilde{q}^{\alpha\zeta}\right]e^{-\theta\left[\tilde{q}^{-\zeta}-q^{-\zeta}\right]}\frac{d\tilde{q}}{\tilde{q}}wL \\
= \; & \left[\frac{1}{\varepsilon-1}+\left(1-\alpha\right)\right]\left(q/Q\right)^{\varepsilon-1}wL+\left[\beta\frac{1}{1-\alpha}\frac{1}{\zeta}+\alpha\right]\frac{m}{\theta}q^{\alpha\zeta}wL \\
& -\beta\int_0^q\left[\left(\tilde{q}/Q\right)^{\varepsilon-1}+\frac{\alpha}{1-\alpha}\frac{m}{\theta}\tilde{q}^{\alpha\zeta}\right]e^{-\theta\left[\tilde{q}^{-\zeta}-q^{-\zeta}\right]}\frac{d\tilde{q}}{\tilde{q}}wL
\end{aligned}
$$

D Evidence from Producer-Level data

One prediction of the model is that if α is higher then, all else equal, the size distribution will have a weakly thicker right tail. This section provides some preliminary evidence on the relationship between intermediate input shares and right tails of size distributions. In particular, I ask whether industries with larger intermediate input shares have size distributions with thicker tails. Before proceeding, it is worth sounding a few notes of caution. First, if suppliers have non-zero bargaining power, intermediate input shares are not the same as α, and differences across industries in the dispersion of marginal cost may confound the comparison. Second, the model abstracts from two features of the reality: producers in the real world use more than a single input and there is a lot of variation within industries in intermediate input shares. Third, treating each industry as a separate economy may be problematic. Finally, it is far from obvious what would serve as the best empirical analogue of an entrepreneur in the model; a case can be made for firms, products, or establishments, although each case has its deficiencies. With those caveats firmly in mind, I proceed to the cross-industry comparison.

I first study French firms. Di Giovanni et al. (2011) compute power law exponents for the tails of the distribution of revenue among French firms in each industry in 2006 and report these in their Table A2. I compare these to intermediate input shares from the French Input Output Tables in 2006 taken from the World Input Output Database (Timmer et al. (2015)). Figure 5 plots each while Table 1 shows regressions of the log of the tail index on the log of the intermediate input share. Di Giovanni et al. (2011) separate industries into tradeable and non-tradeable, and report tail coefficients for the distribution of both domestic and total sales. In line with the theory, all four columns of Table 1 indicate that industries with higher intermediate input shares tend to have thicker right tails (lower exponent). This holds for the distributions of both total and domestic sales, and both across all industries and within categories.

VARIABLES	(1) All Sales	(2) Domestic Sales	(3) All Sales	(4) Domestic Sales
log(Intermediate Input Share)	-0.297***	-0.234**	-0.253**	-0.197
	(0.0986)	(0.0973)	(0.123)	(0.122)
Observations	33	33	33	33
R-squared	0.200	0.135	0.210	0.143
Tradeable FE	No	No	Yes	Yes

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1: French Firms

Note: Robust standard errors are in parentheses. In each regression the log of the tail index is regressed on the log of the intermediate input share. The first and third column use the tail index from the distribution of total revenue, whereas the second and fourth use domestic revenue. The third and fourth use fixed effects indicating whether the industry produces a tradeable good.

Second I study US establishments. I compute power law tail exponents for establishment size distributions for each industry-year in the US using an extract from the Statistics of US Businesses made available by Rossi-Hansberg and Wright (2007) that corresponds to the distribution of employment across establishments in the years 1990, 1992, 1994, 1995, and 1997.[58] The extracts report the number of firms and number of establishments in 45 size categories for various 2-digit SIC industries. Within each industry-year, I compute for each size category the fraction of firms/establishments above the minimum size of the category. I thus have 45 points of the counter-cumulative size distribution. To compute tail coefficients, I choose a size cutoff. I then regress the log of the counter-cumulative distribution on the log of the employment level among employment levels above the cutoff. The baseline results use a size cutoff of 150 employees, but results are similar when using cutoffs of 100, 200, or 400 employees.

Table 2 and Figure 5b show the relationship between the tail coefficients and intermediate input shares from the same industry-year from the BEA. The results are suggestive but less than conclusive.[59] While the point estimates indicate that industries with higher intermediate input

[58] Rossi-Hansberg and Wright (2007) also report an extract for 2000, but I omit these because of the switch from SIC to NAICS.

[59] The analogous figure for US establishments is very similar to Figure 5b.

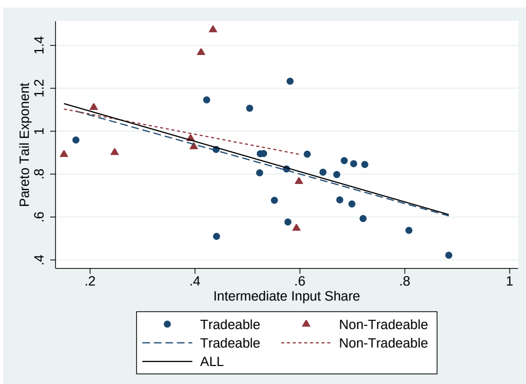
shares have both firm and establishment size distributions with thicker tails, none of the estimates are precise enough to distinguish statistically from zero.

VARIABLES	(1) Firms All	(2) Firms Within Sector	(3) Establishments All	(4) Establishments Within Sector
log(Intermediate Input Share)	-0.144	-0.265	-0.122	-0.235
	(0.197)	(0.218)	(0.104)	(0.151)
Observations	280	280	280	280
R-squared	0.019	0.393	0.020	0.114
Sector FE	NO	YES	NO	YES

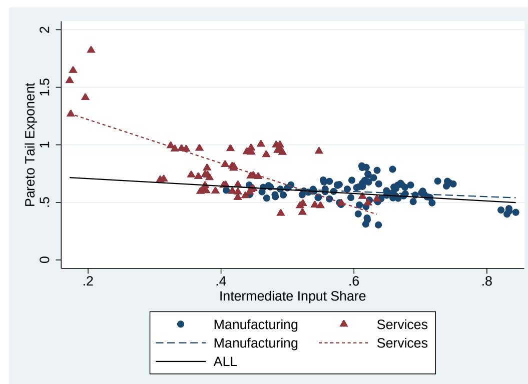
*** p<0.01, ** p<0.05, * p<0.1

Table 2: US Firms and Establishments

Note: Robust standard errors clustered by industry are in parentheses. In each regression the log of the tail index is regressed on the log of the intermediate input share. The second and fourth columns use fixed effects for broad SIC sector.



(a) French Firms



(b) US Firms

Figure 5: Distribution of Customers

For French and US, Figure 5 shows the estimated power law tail exponent for each industry plotted against its intermediate input share. The figures also plot separate regression lines for all industries and for industries within each broad sector (tradeable vs non-tradeable for French firms, manufacturing vs. services for US firms).