How to Build a Production Function*

Sephorah Mangin†

November 2015

Abstract

This paper presents a general recipe for constructing an aggregate production function from a primitive underlying distribution of techniques. The recipe is simple enough to be used as a practical tool. Both the Cobb-Douglas and general C.E.S. production functions are special cases. We uncover an elegant link between the aggregate elasticity of substitution and the virtual valuation function of the underlying distribution of techniques. This delivers an elementary proof that for any well-behaved distribution, the value of the aggregate elasticity of substitution always converges to one in the limit as the capital-labor ratio becomes large. Finally, we show that the asymptotic value of capital's income share is equal to the extreme value *tail index* of the underlying distribution. If capital's share is not asymptotically zero, this implies that the underlying distribution must be fat-tailed, e.g. Pareto (power law) or Fréchet. *JEL* Codes: E23, E25

*I would like to thank Andrew Clausen, Jeff LaFrance, Hongyi Li, John Moore, Matt Rognlie, and Glen Weyl for very helpful comments and conversations.

†Department of Economics, Monash University. Mailing address: PO Box 197 Caulfield East VIC 3145 Australia. Email: sephorah.mangin@monash.edu. Phone: +61 3 9903 2384

1 Introduction

Macroeconomists often use an aggregate production function to describe how the output of an economy is determined by the inputs of capital and labor. The properties of this function are important because they influence the process of economic growth. But where do aggregate production functions come from? Is it possible to construct or *build* such functions, and characterize the features that emerge, instead of simply assuming their properties?

Since the classic result of Houthakker (1955), it has been widely known that the Cobb-Douglas aggregate production function can arise from aggregation across micro-level production units when the underlying distribution of efficiency levels or techniques is Pareto or power law. More recently, Jones (2005) and Lagos (2006) use related approaches to generate Cobb-Douglas aggregate production functions using Pareto distributions.

Power laws have a deep significance in economics and other sciences, and the Pareto distribution emerges naturally in a variety of different contexts.[1] It is therefore perhaps not surprising that the Pareto distribution and the well-known Cobb-Douglas function appear to be two sides of the same coin. However, we might nonetheless wonder: precisely *which features* of a distribution give rise to a production function with specific desired properties?

This paper provides a general recipe for constructing an aggregate production function from *any* underlying distribution of efficiency levels or techniques. The recipe is simple enough to be used as a practical tool. We show exactly how to build or design an aggregate production function with specific properties by providing explicit conditions – on the underlying distribution – under which the standard neoclassical properties and Inada conditions hold.

The general recipe can be explicitly microfounded using extreme value theory by appealing to a powerful result from Gabaix, Laibson, Li, Li, Resnick, and de Vries (2015a). In the present environment, this result implies that the aggregate production function must assume a particular form. We recover the well-known relationship between the Pareto distribution and the Cobb-Douglas

[1] See Gabaix (2009, 2015) for an overview of the many applications of power laws in economics. Gabaix (1999) shows how Zipf's law for cities can arise from proportional random growth with small frictions. Luttmer (2007) and Rossi-Hansberg and Wright (2007) show how the Pareto distribution can emerge as the size distribution of firms.

production function as a canonical special case of this general mapping. We also show that the entire class of constant-elasticity-of-substitution (C.E.S.) aggregate production functions can be derived as special cases. The simple recipe thereby provides a way of unifying the existing results of Houthakker (1955) and later Levhari (1968), as well as Jones (2005) and Lagos (2006).

The elasticity of substitution between capital and labor is an important feature of the aggregate production function. In particular, labor's share will tend to decrease over time with capital accumulation if this elasticity is above one – a fact emphasized by Karabarbounis and Neiman (2014) and Piketty and Zucman (2014) – but it will increase over time if this elasticity is below one. There is currently an active debate in macroeconomics about the value of this elasticity.[2]

We show that the aggregate elasticity of substitution is always less than or equal to one whenever the underlying distribution of techniques is *well-behaved*. The class of such distributions includes almost all standard distributions. We also uncover an elegant link between this macro elasticity and the *virtual valuation function* – a concept familiar to micro theorists since Myerson (1981) first introduced it in the context of optimal auction design. This delivers an elementary proof that, for any well-behaved distribution, the elasticity of substitution always converges to one as the capital-labor ratio becomes large.

The behavior of factor income shares is influenced by properties of the underlying distribution of techniques. For example, if the distribution is Pareto with power-law exponent $\zeta \simeq 3$, capital's share is constant and equal to approximately one third. In general, for *any* well-behaved distribution, the asymptotic value of capital's share – as the capital-labor ratio becomes large – is equal to the *tail index* γ of the underlying distribution, which determines its extreme value domain of attraction.[3] If the limiting value of capital's share is non-zero, this implies that the underlying distribution must be fat-tailed (power tails), i.e. it must be in the Fréchet domain of attraction. This result lends further support to the widespread usage of such distributions, including the Pareto and Fréchet, for modelling firm-level heterogeneity in both trade and macroeconomics.

---

[2] See Antras (2004), Chirinko (2008), Oberfield and Raval (2014), and Rognlie (2014).

[3] Confusingly, the extreme value *tail index* γ of the Pareto distribution is $1/\zeta$ if ζ is the *power-law exponent*, e.g. if $G(x) = 1 - x^{-\zeta}$. A *higher* value of γ (lower ζ) implies fatter tails.

Related Literature. The basic idea of using an underlying distribution of micro-level productivities or techniques to generate an aggregate production function traces its origins back to Houthakker (1955), who first used the Pareto distribution to derive a Cobb-Douglas aggregate production function. Levhari (1968) used a Beta distribution to derive a C.E.S. production function with an elasticity of substitution below one. Recent papers in the spirit of this approach include Jones (2005) and Lagos (2006), who also use Pareto distributions to generate Cobb-Douglas aggregate production functions in different environments.

This paper is perhaps most closely related in spirit to Jones (2005). Jones considers a large number of production units that use local Leontief production technologies. In Jones' paper, ideas represent different ways of combining capital and labor to produce output. The convex hull of all available ideas or techniques across the economy is used to derive a "global" production function. By assuming that the underlying distribution of ideas is Pareto, Jones shows that the global production function is asymptotically Cobb-Douglas in the long run as the total quantity of ideas grows over time.[4]

This paper's approach to modelling ideas or techniques is different to Jones'. Here, techniques are simply the efficiency levels or productivities at which workers can produce output. This approach delivers both simplicity and generality. In particular, we provide a microfounded recipe for constructing an aggregate production function from *any* underlying distribution of techniques that satisfies minimal requirements. This recipe is sufficiently general that it nests the classic results of Houthakker (1955) and Levhari (1968) respectively, as well as Jones (2005) and Lagos (2006). At the same time, the recipe is simple enough that it can readily be used as a practical tool in economic applications.

We appeal to a very general mathematical result found in Gabaix et al. (2015a) to justify the particular form of this recipe. While Gabaix et al. considers equilibrium markups in random demand models with consumer noise, this paper's focus is on aggregate production functions. In particular, we examine how the characteristics of the underlying distribution of techniques influence the properties of the resulting aggregate production function, such as the elasticity of substitution and the behavior of factor income shares. Importantly, the aggregate production function is interpreted as a function of *two* inputs: capital and labor.

[4] See Growiec (2008) for a generalization of some of Jones' aggregation results.

Earlier work using extreme value theory in trade and macroeconomics includes Kortum (1997) and Eaton and Kortum (1999, 2002). Gabaix and Landier (2008) uses some results found in Gabaix et al. (2015a) to study the upper tail of the CEO talent distribution. A number of recent papers highlight the importance of the tail index γ of power law distributions in various applications. Barro and Jin (2011) shows that the size distribution of macroeconomic disasters can be characterized by a power law, with a higher tail index γ (fatter tails) implying a greater equity premium. Gabaix (2011) shows that idiosyncratic firm-level shocks can generate aggregate fluctuations when the firm size distribution has a sufficiently high tail index.[5] Jones and Kim (2015) develops a theory of top income inequality that endogenizes the tail index of the Pareto income distribution. Gabaix, Lasry, Lions, and Moll (2015b) studies the dynamics of both income and wealth inequality and also endogenizes the Pareto tail index.[6]

In a related paper Mangin (2015), I present microfoundations for a class of *unified* aggregate production and matching functions that incorporate labor market frictions. In that paper, wage determination is explicitly modelled using auctions with reservation wages. Factor income shares are affected by the presence of labor market frictions, which leads to unemployment for some workers. In the present paper, by contrast, we abstract entirely from labor market frictions. The focus here is purely on aggregate production functions and their underlying distributions in an otherwise standard neoclassical setting.

Outline. The remainder of this paper is structured as follows. Section 2 provides a microfoundation for the particular recipe we adopt. Section 3 presents results regarding the properties of the aggregate production functions. Section 4 connects some of the results to extreme value theory and derives the recipe formally. Section 5 shows how to extend the domain of the aggregate production function as far as possible. Section 6 presents some key examples, Section 7 discusses the C.E.S. class of aggregate production functions, and Section 8 concludes. All omitted proofs are in the Appendix.

[5] In related work, Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012) shows that idiosyncratic firm-level shocks can generate aggregate fluctuations when the distribution of degrees in the intersectoral network is sufficiently heavy-tailed.

[6] In both papers, "power-law inequality" or "top inequality" is defined as $1/\zeta$ where ζ is the power-law exponent. So *power-law inequality* equals the *tail index* γ here.

2 A Simple Microfoundation

Since the classic results of Houthakker (1955), and later Jones (2005) and Lagos (2006), the close relationship between the Cobb-Douglas function and the Pareto distribution has been well-known. It may therefore be useful to start by considering the simplest such relationship as a purely *mathematical* one.

Specifically, if $G(y) = 1 - y^{-1/\alpha}$ and $f(k) = k^\alpha$ where $y = f(k)$ and $k = K/L$, what exactly is the relationship between the function f and the distribution G? Substituting $y = k^\alpha$ into $G(y)$ and then applying G^{-1} to both sides, we obtain:

$$f(k) = G^{-1}\left(1 - \frac{1}{k}\right). \tag{1}$$

This gives us a preview or a *hint* as to what kind of form we might expect to see for the general mapping between aggregate production functions and their underlying distributions. However, this is a purely mathematical relationship. A deeper question is, how could this relationship possibly arise? Is there a natural way of *microfounding* a general relationship of this form?

Microfoundation. Consider the following simple metaphor. Imagine that God takes a large "bucket" of units of capital that vary with respect to productivity or efficiency. Now imagine that He throws it over His shoulder, so that the units of capital fall randomly across workers.[7] Suppose that workers produce output using the most productive unit of capital that arrives. As the number of units of capital per worker becomes large, what form does the aggregate production function take? What is aggregate output as a function of the inputs?

More precisely, suppose there are $L \in \mathbb{N}$ workers and $K \in \mathbb{N}$ managers (or "entrepreneurs") with a single unit of capital (or a "machine"). Define $k \equiv K/L$, the capital-labor ratio. For simplicity, assume that $k \in \mathbb{N}$.

Each manager has a different *technique* (or "efficiency") drawn from an underlying *distribution of techniques*, G. A manager with technique y can produce y units of the final good by combining one unit of capital with one worker's labor.

Each worker can produce with any one of exactly k managers that are ran-

[7] I thank John Moore for this metaphor.

domly assigned to that worker.[8] We assume that each worker produces output with the manager who has the *best* technique. The expected output of a worker in this economy is therefore $E_G(M_k)$ where M_k is a random variable corresponding to the maximum of k draws from the distribution G.

Now consider the limit as k becomes large. Under mild conditions on the distribution G, we can apply a powerful result from Gabaix et al. (2015a) to obtain a simple expression for expected output per worker. In Section 4, we formally derive the following asymptotic result. As $k \to \infty$, we have

$$E_G(M_k) \propto G^{-1}\left(1 - \frac{1}{k}\right). \tag{2}$$

This result provides at least a partial microfoundation for the aggregate production function we consider. It is "partial" in the following sense. While we considered the limit as $k \to \infty$ to derive the general form (2), we will now consider the properties of the function f_G defined by $f_G(k) \equiv G^{-1}\left(1 - \frac{1}{k}\right)$ as a function of *finite* values of k. Also, while we initially assumed that $k \in \mathbb{N}$, we can now simply extend the domain of f_G to real numbers $k \in \mathbb{R}^+$. Since total output is given by $Y = f_G(k)L$, we will interpret f_G as the intensive form of a constant-returns-to-scale aggregate production function F_G defined by $F_G(K, L) \equiv f_G(k)L$.

In Gabaix et al. (2015a), expression (2) is the asymptotic value of the total *surplus* in a single auction. Here, we are interpreting the average "surplus" $E_G(M_k)$ as representing the average quantity of output per worker. Total output is produced by combining *two* different factors of production: capital K and labor L. The aggregate surplus is actually *created* through the combination of L workers and K managers (or "machines") to generate output.

Remarkably, our construction of the function F_G using a distribution of techniques G appears to make sense as an aggregate production function for a wide range of distributions. Despite the different underlying microfoundations, this simple recipe nests existing results regarding both the Cobb-Douglas and general C.E.S. class of aggregate production functions by Houthakker (1955) and Levhari (1968), as well as Jones (2005) and Lagos (2006).

[8] Alternatively, we could follow the "bucket" metaphor more closely and distribute capital randomly across workers. This would give rise to a Poisson distribution for the number of managers n arriving at each worker, where $n \in \mathbb{N}$ and $E(n) = k$ for $k \in \mathbb{R}^+$. The limiting results are identical and we therefore follow the simpler story.

3 Aggregate Production Function

Now that we have established the general form of the aggregate production function through an elementary microfoundation, we can simply "throw away the ladder."[9] The question of whether the resulting function exhibits the specific properties typically required of an aggregate production function is a mathematical one, while the question of precisely which function best fits the data is an empirical one. We now consider the former question.

Assumption 1. *The distribution of techniques is twice differentiable with cdf* G *and pdf* $g = G' > 0$*, and support* $[y_0, \bar{y})$ *where* $y_0 \in \mathbb{R}^+$ *and* $\bar{y} \in \mathbb{R}^+ \cup \{+\infty\}$.

Let A_1 be the class of all distributions G that satisfy Assumption 1. Given any distribution $G \in A_1$, the aggregate production function is defined as follows.

Definition 1. *For any* $G \in A_1$*, the function* $f_G : [1, \infty) \to \mathbb{R}^+$ *is defined by*

$$f_G(k) \equiv G^{-1}\left(1 - \frac{1}{k}\right) \tag{3}$$

and $F_G : S \to \mathbb{R}^+$ *is defined by* $F_G(K, L) \equiv f_G(k)L$ *where* $k \equiv K/L$.

We use the convention that $G^{-1}(0) = y_0$, so $f_G(1) = y_0$. The intensive production function f_G is defined only for $k \in [1, \infty)$ and the aggregate production function F_G is defined only for $S = \{(K, L) \in \mathbb{R}^+ \times \mathbb{R}^+ | K/L \in [1, \infty)\}$.[10]

We first determine some basic properties that hold for *any* function f_G that is generated by recipe (3) using an underlying distribution of techniques $G \in A_1$.

Lemma 1. *For any* $G \in A_1$*,* f_G *is positive* ($f_G \geq 0$)*, strictly increasing* ($f'_G > 0$)*, twice differentiable, and* $\lim_{k\to\infty} f_G(k) = \bar{y}$.

Letting $y = f_G(k)$, we record the following for future reference. In terms of the underlying distribution G, the marginal product of capital $\frac{\partial F_G(K,L)}{\partial K}$ is

$$f'_G(k) = \frac{(1 - G(y))^2}{g(y)} \tag{4}$$

[9] Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, 6.54.

[10] In Section 5, we discuss how the domain of f_G can be extended to $\mathbb{R}^+$.

and the average product of capital is

$$\frac{f_G(k)}{k} = y(1 - G(y). \tag{5}$$

Lemma 2. *For any $G \in A_1$, if G has a finite mean then* $\lim_{k\to\infty} f'_G(k) = 0$.

Now let $\psi_G : [y_0, \bar{y}) \to \mathbb{R}$ be the *virtual valuation function* of G, defined by

$$\psi_G(y) \equiv y - \frac{1 - G(y)}{g(y)}. \tag{6}$$

The virtual valuation function was first introduced in the context of auction theory in Myerson (1981). An intuitive interpretation of virtual valuations as "marginal revenues" can be found in Bulow and Roberts (1989).[11]

In the present setting, the virtual valuation function represents *wages,* $w_G(k)$. To see this, using (6) and (4), we have

$$\psi_G(y) = f_G(k) - f'_G(k)k. \tag{7}$$

By Euler's theorem, $f_G(k) = f'_G(k)k + \frac{\partial F_G(K,L)}{\partial L}$, so $\psi_G(y)$ equals the marginal product of labor, $\frac{\partial F_G(K,L)}{\partial L}$. If we simply assume that factors are paid their marginal products, we have $w_G(k) = \psi_G(y)$.

Regularity. We say that a distribution G is *regular* on $X \subseteq \mathbb{R}$ if and only if it has a strictly increasing virtual valuation function on X, i.e. $\psi'_G(y) > 0$ for all $y \in X$. The next result states that the function f_G is strictly concave on $f_G^{-1}(X)$ if and only if the underlying distribution G is regular on X. A direct analogue of this fact has been well-known in the auction theory literature since Myerson (1981). We present it here as Lemma 3 for the sake of completeness.

Lemma 3. *For any $G \in A_1$, f_G is strictly concave ($f''_G < 0$) on $f_G^{-1}(X) \subseteq [1, \infty)$ if and only if G is regular on $X \subseteq [y_0, \bar{y})$.*

[11] Bulow and Roberts (1989) highlight an analogy beween the optimal auction design problem and the problem facing a price-discriminating monopolist. The authors interpret the probability that a buyer's valuation is less than the price p as the quantity q demanded. That is, $q \equiv 1 - G(p)$ where G is the distribution of buyers' valuations. The inverse demand function is therefore $p = G^{-1}(1-q)$, the sellers' revenue function is $R(q) \equiv qG^{-1}(1-q)$, and marginal revenue is $R'(q) = \psi(q)$. *Regularity* of G ensures the existence of a unique revenue-maximizing q^* (and p^*).

Proof. Differentiating (6), we obtain

$$\psi_G'(y) = 2 + \frac{g'(y)(1 - G(y))}{g(y)^2}. \tag{8}$$

Differentiating (4) and simplifying using (8), we obtain

$$f_G''(k) = \frac{-(1 - G(y))^3}{g(y)} \psi_G'(y). \tag{9}$$

Clearly, $f_G''(k) < 0$ for all $k \in f_G^{-1}(X)$ if and only if $\psi_G'(y) > 0$ for all $y \in X$. □

Increasing Generalized Hazard Rate. We start with a preliminary definition. Let $\varepsilon_G : [y_0, \bar{y}) \to \mathbb{R}^+$ be the elasticity of $1 - G$ with respect to y,

$$\varepsilon_G(y) \equiv \frac{yg(y)}{1 - G(y)}. \tag{10}$$

This elasticity is called the *generalized hazard rate* since $\varepsilon_G(y) = yh_G(y)$ where $h_G(y) \equiv g(y)/(1 - G(y))$, the *hazard rate*.[12] A distribution G has a weakly increasing generalized hazard rate on $X \subseteq \mathbb{R}$ if and only if $\varepsilon_G'(y) \geq 0$ for all $y \in X$. We call this the *increasing generalized hazard rate* (IGHR) condition.[13]

The IGHR condition is a very mild condition. It is strictly weaker than the increasing hazard rate (IHR) condition since any distribution with an increasing hazard rate automatically satisfies it, while the Pareto distribution is an example of a distribution that has a *decreasing* hazard rate but still satisfies IGHR. In fact, the Pareto distribution is a kind of *boundary* case since it is the unique distribution such that $\varepsilon_G'(y) = 0$ for all y.[14]

Virtually all standard distributions satisfy the IGHR condition. Banciu and Mirchandani (2013) contains an extensive list of distributions that satisfy this condition. Some examples include the Uniform, Exponential, Normal, Logistic, Laplace, Gumbel, Weibull, Gamma, Beta, Pareto, Chi, Lognormal, Cauchy, and

[12] The elasticity $\varepsilon_G(y)$ is sometimes called the *log hazard rate*.

[13] In price theory, this assumption is known as "Marshall's second law of demand" if G is the distribution of consumers' willingness-to-pay. In the operations research literature, the assumption is known as the increasing generalized *failure* rate (IGFR) condition.

[14] The IGHR condition is also strictly weaker than log-concavity, since log-concavity of the density function g implies the IHR condition. See Bagnoli and Bergstrom (2005) for a comprehensive overview of log-concavity.

F distributions. In fact, it is very difficult to find standard distributions that do *not* satisfy this condition for all parameter values.[15]

Well-behaved Distributions. What are the minimal properties of a distribution that ensure recipe (3) yields a sensible aggregate production function? First, we require $G \in A_1$ to ensure that f is positive, strictly increasing, and twice differentiable by Lemma 1. Second, we would like G to have unbounded upper support so that $\lim_{k\to\infty} f(k) = \infty$ by Lemma 1. Next, we require that G has a finite mean to ensure that $\lim_{k\to\infty} f'(k) = 0$ by Lemma 2.

In addition, we would like to find an unbounded range of values y with a weakly positive virtual valuation, i.e. some $y^* \geq y_0$ such that $\psi_G(y) \geq 0$ for all $y \geq y^*$, to ensure that wages are positive. Ideally, we would also like regularity to hold for this range of values to ensure strict concavity of the production function by Lemma 3. It turns out that the IGHR condition delivers what we need.

Definition 2. *We say that G is* well-behaved *if and only if (i) $G \in A_1$; (ii) G has support $[y_0, \infty)$; (iii) G has a finite mean; and (iv) G satisfies IGHR.*

Lemma 4 states that for any well-behaved distribution, there always exists a range of values $[y^*, \infty) \subseteq [y_0, \infty)$ such that $\psi_G(y) \geq 0$ and G is regular on that range. We can therefore simply restrict the domain of f_G to $[k^*, \infty)$ where $k^* = f_G^{-1}(y^*)$ to ensure that both $f''_G(k) < 0$ and $w_G(k) \geq 0$ for all $k \in [k^*, \infty)$. For some distributions, such as the Pareto, we have $y^* = y_0$ and hence $k^* = 1$.

Lemma 4. *If G is well-behaved, there exists $y^* \in [y_0, \infty)$ such that (i) $\psi_G(y) \geq 0$ for all $y \in [y^*, \infty)$; and (ii) G is regular on $[y^*, \infty)$.*

Factor Income Shares. If factors are paid their marginal products, capital's income share $s_K(k)$ equals the output elasticity of f_G with respect to k,

$$s_K(k) = \frac{f'_G(k)k}{f_G(k)}. \tag{11}$$

[15] Banciu and Mirchandani (2013) prove the IGHR class is closed under truncation, but show that the Lévy and Burr distributions do not satisfy the condition for *all* parameter values.

Substituting $y = f_G(k)$, $1 - G(y) = 1/k$, and (4) into (10) yields

$$\varepsilon_G(y)^{-1} = \frac{f'_G(k)k}{f_G(k)}, \tag{12}$$

Combining (11) and (12), we have

$$s_K(k) = \varepsilon_G(y)^{-1}. \tag{13}$$

Labor's share is $s_L(k) = 1 - s_K(k)$. Using (6), (10), and (13), we have

$$s_L(k) = \frac{\psi_G(y)}{y}. \tag{14}$$

Proposition 1 tells us that if the underlying distribution of techniques is well-behaved, factor income shares also behave in a sensible manner.

Proposition 1. *If G is well-behaved, then (i) $s_K(k) \in (0,1]$ for all $k \in [k^*, \infty)$; and (ii) $\lim_{k\to\infty} s_K(k) = \alpha$ for some $\alpha \in [0,1)$.*

Proposition 2 summarizes our results so far regarding the properties of aggregate production functions generated by well-behaved distributions.

Since we restrict f_G to the domain $[k^*, \infty)$ and $F_G(K,L) \equiv f_G(k)L$, we must restrict F_G to the domain $S^* \equiv \{(K,L) \in \mathbb{R}^+ \times \mathbb{R}^+ |\ K/L \in [k^*, \infty)\}$. In Section 5, we will see how to extend the domain of f_G where possible to $\mathbb{R}^+$. We will then be able to consider the remaining neoclassical and Inada conditions, specifically those that apply to the limit $k \to 0$ (i.e. either $K \to 0$ or $L \to \infty$).[16]

Proposition 2. *If G is well-behaved, (i) $F_G : S^* \to \mathbb{R}^+$ is constant-returns-to-scale; (ii) $F_G(K,L) \geq 0$; (iii) F_G is increasing and weakly concave: $\frac{\partial F_G}{\partial K} > 0$, $\frac{\partial F_G}{\partial L} \geq 0$, $\frac{\partial^2 F_G}{\partial K^2} < 0$, $\frac{\partial^2 F_G}{\partial L^2} < 0$; (iv) $\lim_{K\to\infty} F_G(K,L) = +\infty$; (v) $\lim_{K\to\infty} \frac{\partial F_G(K,L)}{\partial K} = 0$; (vi) $\lim_{L\to 0} F_G(K,L) = 0$; and (vii) $\lim_{L\to 0} \frac{\partial F_G(K,L)}{\partial L} = +\infty$.*

3.1 Elasticity of Substitution

We now define the aggregate elasticity of substitution between capital and labor, $\sigma_G : (k^*, \infty) \to \mathbb{R}^+$. Intuitively, the elasticity of substitution represents

[16]Some of the standard properties follow automatically from the others, e.g. see Barro and Sala-i-Martin (2004). However, we prove each explicitly to show which properties of the underlying distribution G are being used and also to be careful regarding the domain of F_G.

the ease of substituting between capital and labor as the relative prices of these inputs change. Since F_G is constant-returns-to-scale, σ_G is a function of k:

$$\sigma_G(k) \equiv -\left(\frac{\partial \log(F_K/F_L)}{\partial \log(K/L)}\right)^{-1} \tag{15}$$

where $F_K = \frac{\partial F_G(K,L)}{\partial K}$ and $F_L = \frac{\partial F_G(K,L)}{\partial L}$.

In general, the elasticity of substitution $\sigma_G(k)$ varies with the capital-labor ratio k. Like other properties of the aggregate production function, the exact nature of this relationship depends upon features of the distribution of techniques.

Lemma 5 identifies an elegant link between the aggregate elasticity of substitution and the virtual valuation function. Specifically, the elasticity of substitution is equal to the *reciprocal* of the elasticity of the virtual valuation function. Since $\psi_G(y) = w_G(k)$ by (7), Lemma 5 is consistent with the result found in Arrow, Chenery, Minhas, and Solow (1961) that the elasticity of substitution equals the elasticity of output per worker with respect to wages.

Given our previous results, the proof of Lemma 5 is simple and it contains an expression for the aggregate elasticity of substitution $\sigma_G(k)$ in terms of the underlying distribution of techniques G. We therefore include the proof here.

Let $\eta_G : (y^*, \bar{y}) \to \mathbb{R}^+$ denote the elasticity of the virtual valuation function,

$$\eta_G(y) \equiv \frac{\psi'_G(y)y}{\psi_G(y)}. \tag{16}$$

Lemma 5. *The elasticity of substitution $\sigma_G : (k^*, \infty) \to \mathbb{R}^+$ satisfies*

$$\sigma_G(k) = \frac{1}{\eta_G(y)} \tag{17}$$

for each $k \in (k^*, \infty)$ and $y = f_G(k)$.

Proof. We start with the following expression from Arrow et al. (1961),

$$\sigma_G(k) = \frac{-f'_G(k)(f_G(k) - kf'_G(k))}{kf_G(k)f''_G(k)}. \tag{18}$$

Substituting (4) and (9) into (18), and setting $f_G(k) = y$ and $k = 1/(1 - G(y))$,

$$\sigma_G(k) = \frac{1 - \frac{1-G(y)}{yg(y)}}{2 + \frac{g'(y)(1-G(y))}{g(y)^2}}. \tag{19}$$

Now, using (6) and (8) together with (16), we have $\sigma_G(k) = 1/\eta_G(y)$. □

It is well-known that if f_G is strictly concave, the elasticity of substitution is less than (greater than) or equal to one if and only if capital's share is weakly decreasing (increasing) in the capital-labor ratio.[17] Combining this well-known result with Lemma 3, and using the fact that capital's share $s_K(k) = \varepsilon_G(y)^{-1}$ by (13), Lemma 6 states if G is regular, the elasticity of substitution $\sigma_G(k)$ is less than (greater than) or equal to one if and only if the generalized hazard rate $\varepsilon_G(y)$ of the underlying distribution G is weakly increasing (decreasing).

Lemma 6. *If G is regular on $X \subseteq [y_0, \infty)$, then for each $y \in X$ and $k = f_G^{-1}(y)$ we have $\sigma_G(k) \leq 1$ if and only if $\varepsilon'_G(y) \geq 0$.*

In the story we told in Section 2 about how the aggregate production function could possibly arise, the "local" elasticity of substitution is zero since the micro-level production technology is essentially Leontief. Given this, one might perhaps think that the aggregate elasticity of substitution must necessarily be less than one for *any* underlying distribution of techniques. In fact, Lemma 6 tells us that it is not the Leontief local production technology *per se* that determines whether the aggregate elasticity of substitution is below one, but rather a specific property of the underlying distribution of techniques.

In particular, Lemma 6 suggests that regular distributions with strictly decreasing generalized hazard rates – if they exist – can generate aggregate production functions with an elasticity of substitution that is always greater than one. Since almost all standard distributions satisfy the IGHR condition, such distributions are difficult to find. In Section 6, we are able to derive such a distribution through reverse engineering – but it does *not* have a finite mean.

Proposition 3 contains two results. The first result states that if G is well-behaved, the elasticity of substitution of the resulting production function is

[17] Indeed, when Hicks (1932) first introduced the concept of the elasticity of substitution, it was in order to characterize the effect of capital accumulation on the factor income distribution.

always less than or equal to one. This follows immediately from Lemmas 4 and 6. The second result states that if G is well-behaved, the value of the elasticity of substitution always converges to one in the limit as the capital-labor ratio becomes large. We include the proof here to highlight its simplicity.

Proposition 3. *If G is well-behaved, then (i) $\sigma_G(k) \in (0,1]$ for all $k \in (k^*, \infty)$; and (ii) $\lim_{k\to\infty} \sigma_G(k) = 1$.*

Proof. The first part follows from Lemmas 4 and 6. For the second part, if G is well-behaved then $\lim_{k\to\infty} s_K(k) = \alpha \in [0,1)$ by Proposition 1. By (14), we have

$$\lim_{y\to\infty} \frac{\psi_G(y)}{y} = 1 - \alpha. \tag{20}$$

Now $\lim_{y\to\infty} \psi_G(y) = \lim_{y\to\infty} s_L(k)y = +\infty$ since $\alpha \neq 1$. By L'Hôpital's rule,

$$\lim_{y\to\infty} \psi_G'(y) = 1 - \alpha. \tag{21}$$

Using (20) and (21) above, we obtain

$$\lim_{y\to\infty} \frac{\psi_G(y)/y}{\psi_G'(y)} = \frac{1-\alpha}{1-\alpha} = 1.$$

So $\lim_{y\to\infty} \eta_G(y) = 1$ and Lemma 5 implies that $\lim_{k\to\infty} \sigma_G(k) = 1$. □

The elementary nature of the proof of Proposition 3 demonstrates how the general mapping (3) allows us to exploit features of the underlying distribution G to establish results about the properties of the aggregate production function F_G that may otherwise be difficult to obtain directly.

There is an existing result in Barelli and de Abreu Pessôa (2003) that proves the Inada conditions imply that the elasticity of substitution is asymptotically equal to one (but only *if* this limit exists).[18] Proposition 3 is consistent with this result. However, the proof here is much simpler than that found in Barelli and de Abreu Pessôa (2003), which first assumes the elasticity of substitution is

[18] Litina and Palivos (2008) identify an intermediate error in Barelli and de Abreu Pessôa (2003) and also highlight the fact that production functions with an elasticity of substitution asymptotically equal to one are *not* necessarily asymptotically Cobb-Douglas, which was mistakenly assumed in the original statement of the result by Barelli and de Abreu Pessôa (2003).

bounded and then establishes that any such production function can be approximated from both above and below by appropriate C.E.S. functions.

4 Relationship to Extreme Value Theory

We say that a distribution G is in the *domain of attraction* of an extreme value distribution G_γ, denoted $G \in D(G_\gamma)$, if there exists a sequence of normalizing constants $a_n > 0$ and b_n such that for each continuity point x, we have $\lim_{n\to\infty} G^n(a_n x + b_n) = G_\gamma(x)$. Intuitively, G_γ is the distribution of the *maximum* of n draws from G as n becomes large. The normalizing constants are necessary to ensure that the limiting distribution makes sense when n goes to infinity.

The parameter γ is called the *extreme value tail index* of G. For brevity, we refer to γ simply as the *tail index* of G.[19] A higher tail index γ implies fatter tails. In particular, the existence of finite moments depends on the value of γ. For $\gamma = 0$, all finite moments exist. For $\gamma > 0$, however, the n-th moment of G exists for $n > 0$, i.e. $E_G(Y^n)$ is finite, if and only if $\gamma < 1/n$.

Definition 3. *We say that G has* tail index γ *if and only if*

$$\lim_{y\to\bar{y}} \frac{d}{dy}\left(\frac{1-G(y)}{g(y)}\right) = \gamma \quad \text{for some } \gamma \in \mathbb{R}. \tag{22}$$

Condition (22) is sometimes called the Von Mises condition. If $G \in A_1$, it is a sufficient condition for $G \in D(G_\gamma)$.[20] Remarkably, if the extreme value distribution G_γ exists, it must be in one of only three classes of distributions: the Fréchet ($\gamma > 0$), Gumbel ($\gamma = 0$), or reverse Weibull class ($\gamma < 0$).

We are now in a position to formally derive expression (2) using a mathematical result from Gabaix et al. (2015a). This result is very general since it incorporates a wide range of distributions with either $\gamma > 0$, $\gamma = 0$, or $\gamma < 0$.

Lemma 7. *If $G \in A_1$ has tail index $\gamma < 1$, in the limit as $k \to \infty$*

$$E_G(M_k) \sim G^{-1}\left(1 - \frac{1}{k}\right)\Gamma(1-\gamma)\ . \tag{23}$$

[19] Confusingly, the "tail index" of a distribution is sometimes defined as $1/\gamma$. Also, γ is sometimes also called the "extreme value index".

[20] To see this, simply apply Theorem 1.1.8, de Haan and Ferreira (2006).

Proof. Since $G \in A_1$ and G has tail index $\gamma < 1$, we can apply Lemma 1 (part 2) and then Theorem 3 from Gabaix et al. (2015a). Using the fact that $G^{-1}(1-\frac{1}{k}) = \bar{G}^{-1}(1/k)$ where $\bar{G}(y) = 1 - G(y)$, we obtain (23) as a special case.[21] □

For simplicity, we normalize expression (23) as follows. Let $M_k' \equiv M_k/\Gamma(1-\gamma)$ where γ is the tail index of G. In the limit as $k \to \infty$, we obtain recipe (3),

$$E_G(M_k') \sim G^{-1}\left(1 - \frac{1}{k}\right). \tag{24}$$

The next proposition tells us that if G is well-behaved, then G has tail index γ and therefore $G \in D(G_\gamma)$ for some $\gamma \in [0, 1)$. This means that Lemma 7 applies to *any* well-behaved distribution. The fact that $\gamma \in [0, 1)$ makes sense: if G is well-behaved, then $\gamma \geq 0$ since G has unbounded upper support; and $\gamma < 1$ since G has a finite mean. So if G is well-behaved, either (*i*) $\gamma \in (0, 1)$ and G is in the domain of attraction of the Fréchet class of distributions (i.e. G has power tails, or is "fat-tailed"); or (*ii*) $\gamma = 0$ and G is in the domain of attraction of the Gumbel class of distributions (i.e. G has exponential tails).

Proposition 4. *If G is well-behaved and $\alpha \equiv \lim_{k\to\infty} s_K(k)$, then G has tail index $\gamma \in [0, 1)$ and $G \in D(G_\gamma)$ where $\gamma = \alpha$.*

Proposition 4 also says that if G is well-behaved, the specific *value* of the tail index γ is equal to the asymptotic value α of capital's income share as the capital-labor ratio becomes large.[22] This implies that if the limiting value of capital's share is non-zero, the underlying distribution must be fat-tailed ($\gamma > 0$), i.e. it must be in the Fréchet domain of attraction. Examples of such distributions include the Pareto or power law distribution and the Fréchet distribution itself.

The Value of Capital's Share. In the U.S., the capital share has historically been relatively stable around the value of *one third*. Indeed, using a Cobb-Douglas

[21] Note that the function *called* G in Gabaix et al. (2015a) is simply the identity function here, i.e. $I(x) = x$, while G here corresponds to the distribution F in that paper.

[22] Mathematically, this can be reconciled with Proposition 2 in Gabaix et al. (2015a) regarding the asymptotic limit-pricing markup when the number of competing firms is large. To see this, observe that capital's income share $s_K(k)$ as the capital-labor ratio k becomes large is identical to μ_n^{LP}/M_n as the number of firms n becomes large, where the markup μ_n^{LP} is defined by $\mu_n^{LP} \equiv M_n - S_n$, and M_n and S_n respectively are the highest and second-highest of n draws.

function and simply setting $\alpha = 1/3$ is quite standard in macroeconomics. But why exactly does the specific value of one third arise?

Proposition 4 identifies an intimate relationship between the distribution of techniques and the behavior of factor shares. In particular, the asymptotic value of capital's share is directly related to the distribution's *tail index,* which governs the existence of finite moments. For any well-behaved distribution G, the n-th moment exists for $n > 0$ if and only if $n < 1/\gamma$ where γ is the tail index of G. Proposition 4 therefore implies that the n-th moment is defined if and only if the asymptotic value α of capital's share is less than $1/n$.

In particular, an asymptotic value α for capital's share that is just below *one third* is the largest value (fattest tails) consistent with a finite mean, variance, and skewness (third moment) of the underlying distribution. This is reminiscent of the intuition highlighted in Gabaix (1999) that Zipf's law arises naturally because a tail index γ just below *one* is the largest value consistent with a finite mean of the Pareto distribution of city sizes. In the present setting, this result suggests there may be some deeper underlying process leading towards the specific value of one third – although we abstract from the details of this process here.[23]

5 Extending the Domain of f_G

One limitation of recipe (3) is that we cannot consider what happens in the limit as $k \to 0$ since $k^* \in [1, \infty)$. However, we can extend the domain of f_G to $\mathbb{R}^+$ where possible. For any distribution $G \in A_1$ with support $[y_0, \infty)$, we have $G(y) = 0$ for all $y \in [0, y_0]$ and $G(y) = \theta(y)$ for all $y \in [y_0, \infty)$ for some function $\theta : [y_0, \infty) \to \mathbb{R}$. If we can extend θ so that $\theta : \mathbb{R}^+ \to \mathbb{R}$, we can extend the function G to $\tilde{G} : \mathbb{R}^+ \to \mathbb{R}$ by simply defining $\tilde{G}(y) \equiv \theta(y)$ for all $y \in \mathbb{R}^+$. (Note that the extended function $\tilde{G}$ may no longer be a cdf.)

We could just define the extended function $f_G : \mathbb{R}^+ \to \mathbb{R}^+$ by $f_G(k) \equiv \tilde{G}^{-1}\left(1 - \frac{1}{k}\right)$ for all $k \in \mathbb{R}^+$. However, we need to restrict the domain of f_G to ensure that $f_G(k) \geq 0$ and $\psi_G(y) \geq 0$ for all k. Note that since $\psi_G(y) = f_G(k) - f'_G(k)k$ and $f'_G(k) > 0$, it is sufficient to ensure that $\psi_G(y) \geq 0$.

[23] In fact, the power-law distribution with $\zeta \simeq 3$ (or $\gamma = 1/3$) is the well-known "cubic" law of stock market returns. Gabaix, Gopikrishnan, Plerou, and Stanley (2003) develops a theory of how this specific distribution can arise in the context of financial markets.

We say that the extended function $\tilde{G}$ is *well-behaved* if G is well-behaved and $\tilde{G} : \mathbb{R}^+ \to \mathbb{R}$ has a weakly increasing generalized hazard rate over its entire domain.[24] We also extend the definition of the virtual valuation function to $\psi_G : \mathbb{R}^+ \to \mathbb{R}$. Analogously to the definition of y^* in the proof of Lemma 4, we define

$$y_{\min} \equiv \inf\{y \in \mathbb{R}^+ | \varepsilon_G(y) \geq 1\}. \tag{25}$$

By similar reasoning to Lemma 4, we have $\psi_G(y) \geq 0$ and $\psi'_G(y) \geq 0$ for all $y \in [y_{\min}, \infty)$. Defining $k_{\min} \equiv f_G^{-1}(y_{\min})$, we can restrict f_G to the extended domain $[k_{\min}, \infty)$ and F_G to the domain $S' \equiv \{(K, L) \in \mathbb{R}^+ \times \mathbb{R}^+ | \ K/L \in [k_{\min}, \infty)\}$.

To consider the remaining properties of F_G as an aggregate production function, we focus on distributions for which $k_{\min} = 0$ and $F_G : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$. Lemma 8 provides sufficient conditions under which $k_{\min} = 0$ and the remaining Inada conditions hold.

Let $\phi_G : \mathbb{R}^+ \to \mathbb{R}^+$ be defined by $\phi_G(y) \equiv y(1 - \tilde{G}(y))$. From (5) above, we know that $\phi_G(y) = f_G(k)/k$, the average product of capital.

Lemma 8. *If $\tilde{G}$ is well-behaved, $\phi'_G(y) \leq 0$ for all $y \in \mathbb{R}^+$, and $\lim_{y\to 0} \phi_G(y) = +\infty$, then (i) $k_{\min} = 0$ and $\lim_{K\to 0} F_G(K, L) = 0$; (ii) $\lim_{K\to 0} \frac{\partial F_G(K,L)}{\partial K} = +\infty$.*

If $\tilde{G}$ is well-behaved, all of the results summarized in Proposition 2 still hold. Proposition 5 states that if, in addition, we have $\phi'_G(y) \leq 0$ for all $y \in \mathbb{R}^+$ and $\lim_{y\to 0} \phi_G(y) = +\infty$, the resulting aggregate production function F_G satisfies *all* of the standard neoclassical properties and Inada conditions.

Proposition 5. *If $\tilde{G}$ is well-behaved, $\phi'_G(y) \leq 0$ for all $y \in \mathbb{R}^+$, and $\lim_{y\to 0} \phi_G(y) = +\infty$, then (i) $F_G : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ is constant-returns-to-scale; (ii) $F_G(K, L) \geq 0$; (iii) For all $K > 0$ and $L > 0$, F_G is strictly increasing and weakly concave: $\frac{\partial F_G}{\partial K} > 0$, $\frac{\partial F_G}{\partial L} > 0$, $\frac{\partial^2 F_G}{\partial K^2} < 0$, $\frac{\partial^2 F_G}{\partial L^2} < 0$; (iv) $\lim_{K\to\infty} F_G(K, L) = \lim_{L\to\infty} F(K, L) = +\infty$; (v) $\lim_{K\to\infty} \frac{\partial F_G(K,L)}{\partial K} = \lim_{L\to\infty} \frac{\partial F_G(K,L)}{\partial L} = 0$; (vi) $\lim_{K\to 0} F_G(K, L) = \lim_{L\to 0} F_G(K, L) = 0$; and (vii) $\lim_{K\to 0} \frac{\partial F_G(K,L)}{\partial K} = \lim_{L\to 0} \frac{\partial F_G(K,L)}{\partial L} = +\infty$.*

[24] We also need to assume that $\tilde{G}$ is twice differentiable and that $\tilde{g} = \tilde{G}' > 0$ in order to satisfy the conditions of Assumption 1.

6 Examples

We might wonder whether the set of underlying distributions G that satisfy Proposition 5 is non-empty. The natural place to start is by considering any function $\phi_G : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\phi'_G(y) \leq 0$ for all $y \in \mathbb{R}^+$, $\lim_{y\to 0} \phi_G(y) = +\infty$, and $\lim_{y\to\infty} \phi_G(y) = 0$. The first two are sufficient conditions in Proposition 5. The latter condition must be true if the distribution G is to be well-behaved, since it must have a finite mean.[25]

The *canonical* function that satisfies all three conditions is $\phi_G(y) = 1/y^p$ where $p > 0$. It remains only to find the corresponding distribution G and check that it is well-behaved. It is easy to see that $G(y) = 1 - y^{-1/\alpha}$ where $\alpha = 1/(p+1)$. Of course, this is just the Pareto distribution with support $[1, \infty)$ and tail index α. It has a finite mean since $\alpha < 1$, and the IGHR condition holds.

By Proposition 5, we conclude that the (extended) aggregate production function F_G defined by recipe (3) satisfies *all* of the standard neoclassical and Inada conditions. As expected, we obtain the familiar Cobb-Douglas production function, $F_G(K, L) = K^\alpha L^{1-\alpha}$. In this sense, the Cobb-Douglas function emerges as the most natural function that satisfies all of the standard neoclassical and Inada conditions. At the same time, the Pareto distribution emerges naturally as a kind of "canonical" underlying distribution in terms of its ability to generate production functions with desirable properties.

The Pareto distribution is a special boundary case with respect to the IGHR condition since it is the *unique* distribution such that $\varepsilon'_G(y) = 0$ for all y. Factor income shares are constant and we have $s_K(k) = \varepsilon_G(y)^{-1} = \alpha$ for all k. For example, if we have $\alpha \simeq 1/3$ then capital's share is roughly one third. Consistent with the results in Section 4 regarding any well-behaved distribution, a value for capital's share just below one third is the largest value (fattest tails) such that the mean, variance, and skewness of G are still finite.

Not all of the conditions listed in Proposition 5 are necessary for all aggregate production functions. The exact requirements of a production function will be determined by the particular context. In many situations, we may not need production functions that are defined for all positive real numbers. Instead, there may be a minimum capital-labor ratio required to produce output. In such

[25] See the Proof of Lemma 2 in the Appendix for details.

cases, the extra conditions identified in Proposition 5 will not be necessary.

One example of an aggregate production function that is defined only for a restricted domain is the following. Let G be the exponential distribution defined by $G(y) = 1 - e^{-y}$. It is straightforward to verify that G is well-behaved and $f_G(k) = \log k$. All of the results in Proposition 2 hold, but Proposition 5 does not apply since we do not have $\phi'_G(y) \leq 0$ for *all* $y \in \mathbb{R}^+$. Instead, we have $k_{\min} = e$ and $f_G : [e, \infty) \to \mathbb{R}^+$. Since G has exponential tails, $\gamma = 0$ and $s_K(k) \to 0$ as $k \to \infty$ by Proposition 4. In addition, we have $\sigma_G(k) \leq 1$ for all k and $\sigma_G(k) \to 1$ in the limit as $k \to \infty$, as we expect from Proposition 3.

7 C.E.S. Class of Production Functions

The constant-elasticity-of-substitution (C.E.S.) aggregate production functions have some peculiar properties. First, they do not satisfy the Inada conditions at both zero and infinity (for any given value of σ). Second, when $\sigma < 1$, output per capita is bounded: $\lim_{k\to\infty} f(k)$ is *finite*. This means that many of the results of previous sections are not applicable. Yet, this class of aggregate production functions is frequently used in macroeconomics. It may therefore be useful to try to construct this class of functions using recipe (3).

For *any* elasticity of substitution $\sigma > 0$, there exists an underlying distribution G_σ that generates the corresponding C.E.S. aggregate production function. For any $\sigma \neq 1$, the distribution G_σ is *not* well-behaved. We obtain the desired distributions through reverse engineering using recipe (3).

Remarkably, the class of distributions G_σ that generates the C.E.S. production functions according to recipe (3) turns out to be essentially the same class as that used by Levhari (1968) and later Lagos (2006) to derive C.E.S. functions. This is despite the fact that the simple recipe used here for constructing such functions is different to the methods used in those papers.

First, we establish a more general result. For any production function f that satisfies certain minimal assumptions, there exists a unique cdf G_f that generates f according to recipe (3). If $G_f \in A_1$ has tail index $\gamma < 1$, we can apply Lemma 7 and hence the construction is consistent with the microfoundations outlined in Sections 2 and 4. Notice that this can include distributions for which $\bar{y} \in \mathbb{R}^+$, i.e. those with *bounded* upper support and tail index $\gamma < 0$.

Proposition 6. *Suppose that $f : [k_{\min}, \infty) \to \mathbb{R}^+$ is a production function that is positive ($f \geq 0$), strictly increasing ($f' > 0$), and twice differentiable. There exists a unique underlying distribution of techniques $G_f \in A_1$ that satisfies (3),*

$$G_f(y) \equiv 1 - \frac{1}{f^{-1}(y)}. \tag{26}$$

The cdf G_f has support $[y_0, \bar{y})$ where $y_0 = f(1)$ and $\bar{y} = \lim_{k\to\infty} f(k)$.[26]

We can now apply Proposition 6. Start with a C.E.S. production function with a constant elasticity of substitution, $\sigma(k) = \sigma \in (0, \infty)$, defined by

$$F(K, L) = \left(\delta K^{\frac{\sigma-1}{\sigma}} + (1-\delta)L^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{27}$$

where $\delta \in (0, 1)$. The intensive form f of this aggregate production function F is

$$f(k) = (\delta k^{\frac{\sigma-1}{\sigma}} + (1-\delta))^{\frac{\sigma}{\sigma-1}} \tag{28}$$

where $k = K/L$ and $f : \mathbb{R}^+ \to \mathbb{R}^+$. To find the unique distribution G_σ that generates this function, we simply use (26) to obtain

$$G_\sigma(y) = 1 - \left(\frac{y^{\frac{\sigma-1}{\sigma}} - (1-\delta)}{\delta}\right)^{-\frac{\sigma}{\sigma-1}}. \tag{29}$$

By Proposition 6, we know that G_σ is a cdf with support $[y_0, \bar{y})$ where $y_0 = f(1) = 1$ and $\bar{y} = \lim_{k\to\infty} f(k)$. The exact nature of the distribution of G_σ depends on the value of the elasticity of substitution σ. On the one hand, if $\sigma < 1$ the distribution has bounded upper support: $\bar{y} = (1-\delta)^{\frac{\sigma}{\sigma-1}}$. On the other hand, if $\sigma \geq 1$ the distribution has unbounded upper support: $\bar{y} = \infty$.

Since f is strictly concave, Lemma 3 tells us that the distribution G_σ defined by (29) is *regular* for any value of σ. However, Lemma 6 implies that G_σ has a strictly *increasing* generalized hazard rate when $\sigma < 1$ and a strictly *decreasing* generalized hazard rate when $\sigma < 1$. In the special case where $\sigma = 1$, Lemma 6 implies that $\varepsilon'_G(y) = 0$ for all y and therefore G_σ has a *constant* generalized hazard rate. As discussed in Section 6, the Pareto distribution is the unique

[26] The solution G_f may not exist in closed form. See Fabinger and Weyl (2015) for detailed conditions under which closed-form solutions exist for related problems in price theory.

distribution with this property, so G_σ must be Pareto if f is Cobb-Douglas.

It is straightforward to show that, for any $\sigma \in \mathbb{R}^+$, the distribution G_σ has tail index $\gamma \leq 1$ for some $\gamma \in \mathbb{R}$. For $\sigma < 1$, the distribution G_σ is in the reverse Weibull domain of attraction ($\gamma < 0$). For $\sigma \geq 1$, the distribution G_σ is in the Fréchet domain of attraction ($\gamma > 0$). The tail index $\gamma(\sigma)$ is given by:

$$\gamma(\sigma) = \begin{cases} 1 & \text{if } \sigma > 1 \\ \delta & \text{if } \sigma = 1 \\ \frac{\sigma-1}{\sigma} & \text{if } \sigma < 1 \end{cases} \tag{30}$$

For any $\sigma \leq 1$, the tail index γ is strictly less than one and hence the distribution G_σ satisfies the assumptions of Lemma 7. However, when $\sigma > 1$, we have $\gamma = 1$ and the result (23) from Gabaix et al. (2015a) does not apply directly.

This example shows that it is indeed possible – at least *mechanically* – to use recipe (26) to find an underlying distribution of techniques that generates an aggregate production function such that $\sigma(k) > 1$ for all k. Crucially, however, the distribution G_σ does *not* have a finite mean when $\sigma > 1$ and Lemma 7 does not apply.[27] It is therefore unclear whether the microfoundations are as compelling in this case where the elasticity of substitution is above one.[28]

For the class of C.E.S. aggregate production functions, at least, an elasticity of substitution less than or equal to one is necessary for the underlying distribution to have a finite mean.[29] If we consider it axiomatic that the distribution of techniques does indeed have a finite mean, this reinforces that $\sigma \leq 1$ is the natural case. This is in addition to the fact that we know already from Lemma 6 that all regular distributions that satisfy the IGHR condition generate aggregate production functions with an elasticity of substitution less than or equal to one. Since this is a very broad class of distributions, the result is widely applicable.

[27] This can be verified directly, but it also follows from the fact that $\gamma = 1$, as well as from Lemma 2 and the fact that $\lim_{k\to\infty} f'(k) \neq 0$ when $\sigma > 1$.

[28] To provide microfoundations, we would need to extend the normalization (24) to include the case $\gamma = 1$. We need to be careful since $\lim_{\gamma\to 1} \Gamma(1-\gamma) = +\infty$. For any G with tail index $\gamma = 1$, consider $\gamma_0 \in [0, 1)$ and $M_k^{'} \equiv M_k/\Gamma(1-\gamma_0)$. We have $\lim_{\gamma_0\to 1} E_G(M_k'; \gamma_0) \sim G^{-1}(1 - 1/k)$ as $k \to \infty$, so we could perhaps simply *define* output per worker $f_G(k)$ using (3) as before.

[29] It is an open question whether there exists some other distribution $G \in A_1$ for which $\varepsilon_G'(y) < 0$ for all y and hence $\sigma_G(k) > 1$ for all k, but which *does* have a finite mean.

8 Conclusion

This paper presents a general, microfounded recipe for constructing an aggregate production function from *any* underlying distribution of techniques. This general recipe nests classic results regarding microfoundations for both the Cobb-Douglas and C.E.S. class of aggregate production functions by Houthakker (1955) and Levhari (1968), and later results by Jones (2005) and Lagos (2006).

Importantly, the recipe is simple enough to be used as a practical tool for economists wishing to construct or design tailored aggregate production functions that feature specific properties. Since the class of standard distributions is extremely rich, and the economist's basic toolkit of aggregate production functions is quite limited, this approach is likely to be fruitful.

We provide explicit conditions under which the standard neoclassical properties and Inada conditions are satisfied. For any well-behaved distribution, the aggregate elasticity of substitution is always less than or equal to one. We identify an elegant link between this macro elasticity and the virtual valuation function that plays an important role in auction theory. This simple connection delivers an elementary proof that the value of the elasticity of substitution must converge to one in the limit as the capital-labor ratio becomes large.

We also prove that, for any well-behaved distribution, the asymptotic value of capital's income share is equal to the extreme value *tail index* of the underlying distribution. In particular, this means that if the limiting value of capital's share is non-zero, the underlying distribution must be fat-tailed. This theoretical result lends further support to the widespread usage of such distributions, including the Pareto and Fréchet distributions, for modelling firm-level heterogeneity of various kinds in both macroeconomics and trade.

Throughout most of this paper, we have treated the distribution of techniques as a *primitive* that is used to construct the aggregate production function. However, it is also possible to work backwards to recover the distribution of techniques for a given industry, sector, or country. We could potentially start with a set of observations (y, k) of output per worker $y = f(k)$ and the capital-labor ratio k and then *estimate* the corresponding distribution G using the inverse mapping. We leave empirical questions of this nature as a topic for future research.

9 Appendix: Omitted Proofs

Proof of Lemma 1

Proof. Clearly, f_G is positive since $[y_0, \bar{y}) \subseteq \mathbb{R}^+$ and f_G is twice differentiable since $G \in A_1$. We have $f_G'(k) = \frac{d}{dk} G^{-1}\left(1 - \frac{1}{k}\right) = \frac{1}{k^2 g\left(G^{-1}\left(1-\frac{1}{k}\right)\right)} > 0$ for all $k \in [1, \infty)$. Finally, $\lim_{k\to\infty} f_G(k) = \lim_{k\to\infty} G^{-1}\left(1 - \frac{1}{k}\right) = G^{-1}(1) = \bar{y}$. □

Proof of Lemma 2

Proof. If $\lim_{k\to\infty} f_G(k) = \bar{y} \in \mathbb{R}^+$ then $\lim_{k\to\infty} f_G'(k) = 0$. If $\lim_{k\to\infty} f_G(k) = \infty$, we have $\lim_{k\to\infty} f_G'(k) = \lim_{k\to\infty} \frac{f_G(k)}{k}$ by L'Hôpital's rule. Using (5), we have $\frac{f_G(k)}{k} = y(1-G(y))$ so $\lim_{k\to\infty} f_G'(k) = \lim_{y\to\infty} y(1-G(y))$. If G has a finite mean, $\int_{y_0}^{\infty} yg(y)dy < \infty$. Since $0 \leq y(1-G(y)) = y\int_y^{\infty} g(t)dt \leq \int_y^{\infty} tg(t)dt < \infty$ for all y and $\lim_{y\to\infty} \int_y^{\infty} tg(t)dt = 0$, we have $\lim_{y\to\infty} y(1-G(y)) = \lim_{k\to\infty} f_G'(k) = 0$. □

Proof of Lemma 4

Proof. For part (i), first observe that $\psi_G(y) > 0$ if and only if $\varepsilon_G(y) > 1$ by (6) and (10). By Theorem 2 from Lariviere (2006), if G satisfies the IGHR condition then $\lim_{y\to\infty} \varepsilon_G(y) > 1$ (including possibly infinite) if and only if G has a finite mean. If G is well-behaved, it has a finite mean so $\lim_{y\to\infty} \varepsilon_G(y) > 1$. Either $\varepsilon_G(y) > 1$ and hence $\psi_G(y) > 0$ for all $y \in [y_0, \infty)$, or alternatively $\varepsilon_G(y) < 1$ for some $y \in [y_0, \infty)$. In the latter case, if G is well-behaved then $\varepsilon_G'(y) \geq 0$ for all y and since $\lim_{y\to\infty} \varepsilon_G(y) > 1$ there exists a non-empty range of values $\varepsilon_G^{-1}(1) \subseteq [y_0, \infty)$ such that $\varepsilon_G(y) = 1$ for all $y \in \varepsilon_G^{-1}(1)$. Now define

$$y^* = \inf\{y \in [y_0, \infty) | \varepsilon_G(y) \geq 1\}. \tag{31}$$

Since $\varepsilon_G(y^*) = 1$, for all $y \in [y^*, \infty)$ we have $\varepsilon_G(y) \geq 1$ and hence $\psi_G(y) \geq 0$. For part (ii), using (6), and (10), we have

$$\psi_G(y) = y\left(1 - \frac{1}{\varepsilon_G(y)}\right). \tag{32}$$

Differentiating (32), we have

$$\psi_G'(y) = 1 - \frac{1}{\varepsilon_G(y)} + \frac{y\varepsilon_G'(y)}{\varepsilon_G(y)^2}. \tag{33}$$

Since $\varepsilon_G(y) \geq 1$ for all $y \in [y^*, \infty)$, we have $1 - \varepsilon_G(y)^{-1} \geq 0$. Since G is well-behaved, we have $\varepsilon_G'(y) \geq 0$, which implies $\psi_G'(y) \geq 0$ for all $y \in [y^*, \infty)$. □

Proof of Proposition 1

Proof. For part (i), since $G(y) < 1$ for all y, we have $\varepsilon_G(y)^{-1} > 0$ for all y and hence $s_K(k) > 0$ for all k. If G is well-behaved, we have $\psi_G(y) \geq 0$ and hence $\varepsilon_G(y)^{-1} \leq 1$ for all $y \in [y^*, \infty)$ by Lemma 4, so $s_K(k) \in (0, 1]$ for all $k \in [k^*, \infty)$. For part (ii), if G is well-behaved, then $\varepsilon_G'(y) \geq 0$ for all $y \in [y^*, \infty)$ and $\varepsilon_G(y)^{-1}$ is weakly decreasing in y. For all $y \in [y^*, \infty)$, we have $\varepsilon_G(y)^{-1} \in (0, 1]$ by part (i) so $\varepsilon_G(y)^{-1}$ is both weakly decreasing in y and bounded below by zero, hence $\lim_{y\to\infty} \varepsilon_G(y)^{-1}$ exists and $\lim_{k\to\infty} s_K(k) \equiv \alpha \in [0, 1]$. By Theorem 2 from Lariviere (2006), $\lim_{y\to\infty} \varepsilon_G(y) > 1$ if and only if G has a finite mean. Since G is well-behaved, it has a finite mean so $\lim_{y\to\infty} \varepsilon_G(y)^{-1} < 1$ and $\alpha \in [0, 1)$. □

Proof of Proposition 2

Proof. (i) By construction, $F_G(\lambda K, \lambda L) = \lambda f_G(k)L$ for all $\lambda > 0$. (ii) Clear. (iii) Follows from Lemmas 1, 3, and 4. It is straightforward to show that $\frac{\partial^2 F_G}{\partial L^2} < 0$ whenever $f''(k) < 0$. Note that since $\psi_G(y^*) = 0$, we have $\frac{\partial F_G}{\partial L} > 0$ for all $k > k^*$. (iv) Follows from Lemma 1. (v) Follows from Lemma 2 since if G is well-behaved it has a finite mean. (vi) We have $\lim_{L\to 0} F_G(K, L) = \lim_{k\to\infty} f_G(k)L$ where $L = K/k$, so $\lim_{L\to 0} F_G(K, L) = \lim_{k\to\infty} \frac{f_G(k)}{k} K = \lim_{k\to\infty} f_G'(k)K$ since $\lim_{k\to\infty} f_G(k) = +\infty$. Therefore $\lim_{L\to 0} F_G(K, L) = 0$ follows from Lemma 2. (vii) First, $\lim_{L\to 0} \frac{\partial F_G(K,L)}{\partial L} = \lim_{y\to\infty} \psi_G(y)$ since $y \to \infty$ as $k = K/L \to \infty$. By (32), $\psi_G(y) = y\,(1 - \varepsilon_G(y)^{-1})$, so $\lim_{L\to 0} \frac{\partial F_G(K,L)}{\partial L} = \lim_{y\to\infty} y\,(1 - \varepsilon_G(y)^{-1})$. By Proposition 1, $\lim_{y\to\infty} \varepsilon_G(y)^{-1} = \alpha \in [0, 1)$, so $\lim_{L\to 0} \frac{\partial F_G(K,L)}{\partial L} = +\infty$. □

Proof of Lemma 6

Proof. Using (11) and (18), it is straightforward to verify that if $f_G''(k) < 0$, then $\sigma_G(k) \leq 1$ if and only if $s_K'(k) \leq 0$. If G is regular on X, then f_G is strictly concave on $f_G^{-1}(X)$ by Lemma 3. Since $s_K(k) = \varepsilon_G(y)^{-1}$ by (13), this implies that $\sigma_G(k) \leq 1$ if and only if $\varepsilon'_G(y) \geq 0$. $\square$

Proof of Proposition 4

Proof. Condition (22) is equivalent to:

$$\lim_{y \to \bar{y}} \frac{g'(y)(1 - G(y))}{g(y)^2} = -\gamma - 1 \text{ for some } \gamma \in \mathbb{R}. \tag{34}$$

Using (8), it is easy to see that condition (34) is equivalent to

$$\lim_{y \to \infty} \psi'_G(y) = 1 - \gamma \text{ for some } \gamma \in \mathbb{R}. \tag{35}$$

By (21), we have $\lim_{y \to \infty} \psi'_G(y) = 1 - \alpha$ where $\alpha \in [0, 1)$. So (35) holds and G has tail index $\gamma = \alpha$. Therefore $G \in D(G_\gamma)$. $\square$

Proof of Lemma 8

Proof. First, notice that $\lim_{y \to 0} \phi_G(y) = +\infty$ implies $\lim_{y \to 0} \tilde{G}(y) = -\infty$. For part (i), for distributions with $\phi'_G(y) \leq 0$ we have $\varepsilon_G(y) \geq 1$ for all $y \in \mathbb{R}^+$ so $y_{\min} = 0$. By definition, $k_{\min} \equiv f_G^{-1}(y_{\min}) = f_G^{-1}(0)$ if $y_{\min} = 0$. So $k_{\min} = 0$ if and only if $f_G^{-1}(0) = 0$, which is true if and only if $\tilde{G}(0) = -\infty$. Next, $\lim_{K \to 0} F_G(K, L) = \lim_{k \to 0} f_G(k)L = y_{\min}L = 0$. For part (ii), using part (i) we have $\lim_{K \to 0} \frac{\partial F_G(K,L)}{\partial K} = +\infty$ if and only if $\lim_{k \to 0} \frac{f_G(k)}{k} = +\infty$, which holds if and only if $\lim_{y \to 0} \phi_G(y) = +\infty$. $\square$

Proof of Proposition 5

Proof. We need only show that $\lim_{L \to \infty} F(K, L) = +\infty$ and $\lim_{L \to \infty} \frac{\partial F_G(K,L)}{\partial L}$. All other results follow immediately from Proposition 2 and Lemma 8. To start with, we have $\lim_{L \to \infty} \frac{\partial F_G(K,L)}{\partial L} = \lim_{y \to 0} \psi_G(y)$ since $\lim_{k \to 0} f_G(k) = 0$ by Lemma 8. Also, by (32), $\psi_G(y) = y\left(1 - \varepsilon_G(y)^{-1}\right)$ so $\lim_{y \to 0} \psi_G(y) = \lim_{y \to 0} y\left(1 - \varepsilon_G(y)^{-1}\right)$. Since $\phi'_G(y) \leq 0$ for all $y \in \mathbb{R}^+$, we have $\varepsilon_G(y) \geq 1$ for all $y \in \mathbb{R}^+$, and $\varepsilon'_G(y) \geq 0$ since $\tilde{G}$ is well-behaved. So $\lim_{y \to 0} \varepsilon_G(y) \in [1, \infty)$ and therefore $\lim_{y \to 0} y\left(1 - \varepsilon_G(y)^{-1}\right) = \lim_{L \to \infty} \frac{\partial F_G(K,L)}{\partial L} = 0$. Next, we have $\lim_{L \to \infty} F_G(K, L) = \lim_{k \to 0} f_G(k)L$, which equals $\lim_{k \to 0} \frac{f_G(k)}{k} K = +\infty$ since $\lim_{y \to 0} \phi_G(y) = +\infty$. $\square$

Proof of Proposition 6

Proof. First, the function G_f has support $[y_0, \bar{y})$ where $y_0 = f(1) \in \mathbb{R}^+$ and $\bar{y} = \lim_{k\to\infty} f(k)$. We have $y_0 = f(1)$ since $G(y) \geq 0$ if and only if $f^{-1}(y) \geq 1$, which is true if and only if $y \geq f(1)$. We also have $\bar{y} = \lim_{k\to\infty} f(k)$ since $\lim_{y\to\bar{y}} G_f(y) = \lim_{y\to\bar{y}} (1 - f^{-1}(\bar{y})^{-1}) = 1$. Since $\lim_{y\to\infty} G_f(y) = 1$, we have $\int_{y_0}^{\bar{y}} g_f(y)dy = 1$ and G_f is a cdf. Next, G_f is clearly twice differentiable since f is. Finally, $g_f(y) = G_f'(y) = \frac{d}{dy}\left(1 - \frac{1}{f^{-1}(y)}\right) = \frac{\frac{d}{dy} f^{-1}(y)}{(f^{-1}(y))^2} = \frac{1}{f'(f^{-1}(y))(f^{-1}(y))^2} > 0$. □

Tail index of C.E.S. class of distributions

Proof. First, the pdf g_σ of the function G_σ is

$$g_\sigma(y) = \frac{y^{-1/\sigma}}{\delta}\left(\frac{y^{\frac{\sigma-1}{\sigma}} - (1-\delta)}{\delta}\right)^{-\frac{\sigma}{\sigma-1}-1}. \tag{36}$$

Next, substituting in (29) and (36) and simplifying, we have

$$\lim_{y\to\bar{y}} \frac{g'(y)(1-G(y))}{g(y)^2} = \lim_{y\to\bar{y}} \frac{(1-\delta)y^{\frac{1-\sigma}{\sigma}}}{\sigma} - 2. \tag{37}$$

Using (34) and (37) and rearranging, the tail index γ satisfies

$$\gamma = \lim_{y\to\bar{y}}\ 1 - \frac{(1-\delta)y^{\frac{1-\sigma}{\sigma}}}{\sigma}. \tag{38}$$

Finally, equation (38) yields $\gamma(\sigma)$ as in equation (30). □

References

D. Acemoglu, V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi. The Network Origins of Aggregate Fluctuations. *Econometrica*, 80(5):1977–2016, 2012.

P. Antras. Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution. *Contributions to Macroeconomics*, 4(1), 2004.

K. J. Arrow, H. B. Chenery, B. S. Minhas, and R. M. Solow. Capital-Labor Substitution and Economic Efficiency. *Review of Economics and Statistics*, 43 (3):225–250, 1961.

M. Bagnoli and T. Bergstrom. Log-concave probability and its applications. *Economic Theory*, 26(2):445–469, 2005.

M. Banciu and P. Mirchandani. Technical note: New results concerning probability distributions with increasing generalized failure rates. *Operations Research*, 61(4):925–931, 2013.

P. Barelli and S. de Abreu Pessôa. Inada conditions imply that production function must be asymptotically cobb-douglas. *Economics Letters*, 81(3):361–363, 2003.

R. J. Barro and T. Jin. On the Size Distribution of Macroeconomic Disasters. *Econometrica*, 79(5):1567–1589, 2011.

R. J. Barro and X. Sala-i-Martin. *Economic Growth, Second Edition*. The MIT Press, Cambridge, Massachusetts, 2004.

J. Bulow and J. Roberts. The Simple Economics of Optimal Auctions. *The Journal of Political Economy*, 97(5):1060–1090, 1989.

R. Chirinko. Sigma: The Long and Short of it. *Journal of Macroeconomics*, 30 (2):671–686, 2008.

L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer, New York, 2006.

J. Eaton and S. Kortum. International technology diffusion: Theory and measurement. *International Economic Review*, 40(3):537–570, 1999.

J. Eaton and S. Kortum. Technology, geography, and trade. *Econometrica*, 70 (5):1741–1779, 2002.

M. Fabinger and E. G. Weyl. A Tractable Approach to Pass-Through Patterns. *Working paper*, 2015.

X. Gabaix. Zipf's Law for Cities: An Explanation. *The Quarterly Journal of Economics*, pages 739–767, 1999.

X. Gabaix. Power Laws in Economics and Finance. *Annual Review of Economics*, 1:255–293, 2009.

X. Gabaix. The Granular Origins of Aggregate Fluctuations. *Econometrica*, 79 (3):733–772, 2011.

X. Gabaix. Power laws in economics: An introduction. *The Journal of Economic Perspectives (forthcoming)*, 2015.

X. Gabaix and A. Landier. Why has CEO pay increased so much? *The Quarterly Journal of Economics*, 123(1):49–100, 2008.

X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley. A theory of power-law distributions in financial market fluctuations. *Nature*, 423(6937):267–270, 2003.

X. Gabaix, D. Laibson, D. Li, H. Li, S. Resnick, and C. G. de Vries. The Impact of Competition on Prices with Numerous Firms. *Working Paper*, 2015a.

X. Gabaix, J.-M. Lasry, P.-L. Lions, and B. Moll. The Dynamics of Inequality. *Working paper*, 2015b.

J. Growiec. Production functions and distributions of unit factor productivities: Uncovering the link. *Economics Letters*, 101(1):87–90, 2008.

J. R. Hicks. The Theory of Wages. 1932.

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

C. I. Jones. The Shape of Production Functions and the Direction of Technical Change. *Quarterly Journal of Economics*, 120(2):517–549, 2005.

C. I. Jones and J. Kim. A Schumpeterian Model of Top Income Inequality. *Working paper*, 2015.

L. Karabarbounis and B. Neiman. The Global Decline of the Labor Share. *Quarterly Journal of Economics*, 129(1):61–103, 2014.

S. Kortum. Research, Patenting, and Technological Change. *Econometrica*, 65 (6):1389–1419, 1997.

R. Lagos. A Model of TFP. *Review of Economic Studies*, 73(4):983–1007, 2006.

M. A. Lariviere. A note on probability distributions with increasing generalized failure rates. *Operations Research*, 54(3):602–604, 2006.

D. Levhari. A Note on Houthakker's Aggregate Production Function in a Multi-firm Industry. *Econometrica*, 36(1):151–154, 1968.

A. Litina and T. Palivos. Do Inada Conditions Imply that Production Functions must be Asymptotically Cobb-Douglas? a comment. *Economics Letters*, 99 (3):498–499, 2008.

E. G. Luttmer. Selection, Growth, and the Size Distribution of Firms. *The Quarterly Journal of Economics*, pages 1103–1144, 2007.

S. Mangin. A theory of production, matching, and distribution. *Working Paper*, 2015.

R. B. Myerson. Optimal auction design. *Mathematics of operations research*, 6 (1):58–73, 1981.

E. Oberfield and D. Raval. Micro data and macro technology. *Working Paper*, 2014.

T. Piketty and G. Zucman. Capital is back: Wealth-income ratios in rich countries 1700-2010. *Quarterly Journal of Economics*, 129(3):1255–1310, 2014.

M. Rognlie. A Note on Piketty and Diminishing Returns to Capital. *Working Paper*, 2014.

E. Rossi-Hansberg and M. L. Wright. Urban Structure and Growth. *The Review of Economic Studies*, 74(2):597–624, 2007.