Automation, Growth and Factor Shares

Joseba Martinez*

New York University

This Version: February 2018

Please Click Here to Download Latest Version

Abstract

This paper investigates the extent to which automation can explain the observed fall in labor's share of income in the United States in the last 30 years. I model the production process as a set of tasks that can be performed by labor or automated machinery (capital). Aggregating over firms that operate capital with differing degrees of automation, total output of the economy is given by a Constant Elasticity of Substitution (CES) function, but with parameters determined endogenously by the distribution of automation technology across firms. This model of the aggregate production function can reconcile three important empirical findings on US production and growth that the canonical CES model cannot: declining labor shares, aggregate capital-labor complementarity, and capital-biased technical progress. Using industry-level data, including a novel measure of aggregate task inputs into production, I find evidence that automation was a significant driving force of the US labor share between 1972-2010.

*I owe a profound debt of gratitude to my advisers Mark Gertler, Virgiliu Midrigan, Thomas Philippon and Thomas Sargent for their invaluable guidance and encouragement. I thank Diego Anzoategui, Jess Benhabib, Jaroslav Borovička, Katarína Borovičková, Daniel Greenwald, Boyan Jovanovic, Ricardo Lagos, Matthias Kehrig, Diego Perez, Venky Venkateswaran and Gianluca Violante for many comments, suggestions and insights. Contact details: joseba.martinez@nyu.edu; www.josebamartinez.me.

1 Introduction

Labor's share of income in the United States has fallen over the last 30 years. This paper develops a theoretical model of the link between automation and the labor share and, through the lens of the model, provides evidence on the extent to which the observed decline might have been caused by automation.

Automation in the model is a form of embodied technical progress that enables the replacement of workers with capital in the completion of production tasks. Firms operate capital with varying degrees of automation: some firms' capital can perform a greater range of production tasks than others. Aggregate output in this economy can be represented using the canonical constant elasticity of substitution (CES) function with capital and labor aggregates as inputs; but with parameters that are themselves functions of the level and distribution of automation technology in the economy. This model of the aggregate production function delivers a sharp characterization of the effect of automation on the labor share: the labor share *decreases* if the *most* automated firms in the economy become more automated; but *increases* if the *least* automated firms become more automated.

To assess the empirical validity of this prediction, I construct a measure of aggregate task inputs into production in the US economy by combining census data from the Current Population Survey with detailed occupation level data from the Occupational Information Network (O*NET). The use of job task data to study labor market outcomes is by now well established, notably in the literature following Autor et al. (2003) that studies the effect of technology on job polarization among other labor market outcomes. Acemoglu and Autor (2011) combine CPS and O*NET data to construct broad occupational/task categories. But to my knowledge this is the first paper that uses this type of data as an input into a macro accounting exercise. I combine the data on task inputs with industry-level income accounts (from KLEMS, Jorgenson et al. (2012)) to measure the effect of automation on US industry labor shares in the period 1972-2010.

The empirical findings suggest an important role for automation in the evolution of the labor share at the industry level, particularly in manufacturing: the fitted model generates more than half of the decline in the manufacturing labor share.

The aggregate elasticity of substitution - σ in the standard notation - between capital and labor is a crucial parameter in macroeconomics and especially in growth theory[1]. In particular, the value of σ determines the response of the labor share to changes in the capital-labor ratio: if $\sigma > 1$, so that capital and labor are gross substitutes, the labor share *decreases* when the capital-labor *increases*, and vice versa for $\sigma < 1$. This is an important relationship because the period over which the labor share has fallen coincides with a period of rapid improvement in the quality of capital goods, which is usually modeled as capital- or investment-biased technical change.

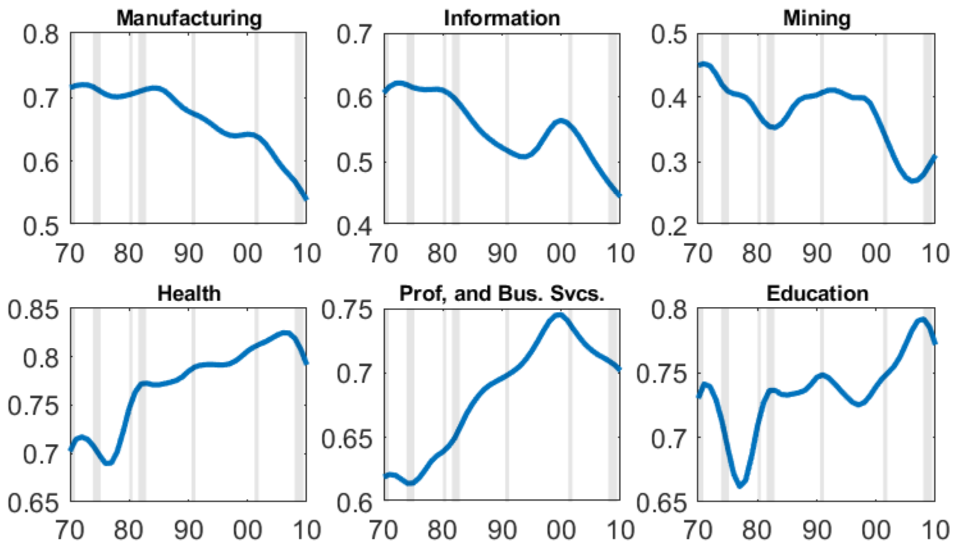
As shown by Karabarbounis and Neiman (2014), if this biased technical change leads to increased capital accumulation and an increase in the capital-labor ratio, $\sigma > 1$ is required to explain the falling labor share. The preponderance of empirical estimates however point to $\sigma < 1$[2]. Oberfield and Raval (2014) show that, since this implies that capital accumulation would increase the labor share, other forces must be driving the observed fall. In the model of aggregate production function developed in Section 2, capital and labor are gross complements. The dynamic model in Section 3 shows how the investment, entry and exit decisions of firms operating capital with differing degrees of automation endogenously determine the parameters of the aggregate production function. Viewed from the perspective of the aggregate production function, capital biased technical change in the form of increased automation lowers the labor share by increasing the capital share parameter α.

The empirical analysis in this paper focuses on industry-level labor shares. Figure 1, which plots shares for a subset of industries, shows that trends in industry labor shares level are both significant and markedly different across sectors. The empirical analysis in this paper uses industry-level data in order to make use of the information contained in this divergence: the conjecture that underpins the exercise is that it reflects, at least in part, the extent to which automation of production tasks has affected industries differentially.

Related literature This paper contributes to a growing literature that explores the link between automation, growth, productivity and factor shares building on task-based production models and studying the effect of automation on the aggregate production function. An early contribution in this vein is Zeira (1998), who starts with the observation that labor replacing embodied technical progress is a long run phenomenon and develops an elegant model

[1] de La Grandville (2016) gives a comprehensive treatment of the importance of σ.

[2] See Chirinko (2008) for a survey of estimates of σ.



Source: US KLEMS data. Labor share is compensation divided by value added. Series are HP filtered (filter parameter 6.25)

Figure 1: Labor shares for 6 industries, 1970-2010

in which task automation (effectively) changes the exponents of the aggregate Cobb-Douglas production function, but also necessarily implies increasing capital shares and accelerating economic growth.

The model in this paper starts from different primitives but has features in common with Acemoglu and Restrepo (2016). The authors develop a comprehensive theory of automation featuring endogenous automation and creation of new tasks that only labor can perform (the "race" to which the title refers); when these two forces are evenly matched, the economy has a balanced growth path. A closely related condition has to hold for the dynamic model described in Section 3 to have a balanced growth path. An implication of their theory is that a long-run negative relationship between capital accumulation and the labor share does not necessarily imply that labor and capital are gross aggregate substitutes, if the allocation of factors to tasks is endogenous. The aggregative model of the production function developed in this paper also has this implication, as discussed above.

Aghion et al. (2017) present a model of automation that builds on the structure of Zeira (1998). The authors examine the ability of such a model to generate dynamics consistent with the Kaldor growth facts. Among other insights, the model connects automation with the emergence of a Baumol (1967) "cost disease"; the authors also explore the effect of a possible technological singularity leading to an explosion in economic growth. Hemous and Olsen (2016) develop a model of automation and horizontal innovation that endogenously links automation to the rise in the skilled-unskilled wage premium. Susskind (2017) develops a model of automation in which labor is eventually completely immiserated due to the development of "advanced capital" that can replace labor in advanced tasks.

This paper makes two main contributions to this literature. First, starting from an explicit micro task production technology, the aggregation procedure I employ generates a canonical aggregate CES production function with parameters that are determined endogenously by the distribution of automation in the economy. This approach facilitates comparison with the extant literature and can help to explain how the canonical model might be able to accommodate features of economic growth that are at apparent odds from the perspective of a model with a fixed aggregate production function. Second, the model in this paper delivers a sharp characterization of the link between automation and the labor share that is simple enough to be taken to data.

The model in this paper draws from a variety of theoretical foundations. The micro task production function that I propose is closely related to that of Becker and Murphy (1992) and to models in which production is organized as a hierarchy, starting with Garicano (2000) and Garicano and Rossi-Hansberg (2006). To my knowledge, Geerolf (2017) is the first paper to make explicit the link between a Garicano (2000)-type production function and the Pareto distribution, an insight which I apply in my aggregation procedure.

The aggregation procedure I implement was first introduced in Houthakker (1955), which shows that aggregation of micro-production units operating a Leontieff technology with Pareto-distributed productivities results in a Cobb-Douglas aggregate production function. Levhari (1968) and Sato (1975) extend the result to general CES functions, in a static setting in which the productivity distribution itself is the primitive. Jones (2005) applies this aggregation result in a growth context to think about how the aggregate production changes over time, and how technological change matters for the shape of the production function. This paper's central concern is similar in spirit, but my focus is on automation specifically. The model in this paper is closest to Lagos (2006), who develops an aggregative model of total factor productivity. The model of Section 2 can be thought of as extending his work to a model of both TFP and the share parameter (α).

The dynamic version of the model features a putty-clay vintage capital structure, first introduced in Johansen (1959). The implementation of the putty-clay technology in this model is similar to Gilchrist and Williams (2000). As in Jovanovic and Yatsenko (2012), and in contrast to many vintage capital models, the model in this paper features investment in both old and new vintages of capital.

This paper explores the link between the US labor share and automation; a burgeoning literature documents recent movements in the US labor share and proposes several explanations for its decline. Elsby et al. (2013) document trends in both aggregate and industry level labor shares in the US, and provide an overview of possible causes for observed declines. Rognlie (2015) shows that the fall in the labor share is largely accounted for by increased rents accruing to housing. Gutierrez (2017), Barkai (2017), Caballero et al. (2017) and De Loecker and Eeckhout (2017) focus on increased rents and market power. Hartman-Glaser et al. (2017), Autor et al. (2017) and Kehrig and Vincent (2017) document that changes in the distribution of industry output towards lower labor share firms are responsible for the overall decline. Koh et al. (2016) suggest that the aggregate decline in the labor share is largely accounted for by increased depreciation of intangible capital. Karabarbounis and Neiman (2014) focus on capital-labor substitution driven by the sustained decline in the relative price of capital; in cross-country regressions they estimate the elasticity of substitution parameter of a CES production function $\sigma > 1$, consistent with their hypothesis.

Oberfield and Raval (2014) show that, given that their estimate of the aggregate elasticity of substitution between labor and capital in manufacturing is smaller than one, increased capital accumulation in response to the declining relative price of capital cannot explain the decrease in the labor share in manufacturing. In fact, they find the opposite to be true: since $\sigma < 1$, a higher capital-labor ratio increases the labor share. They conclude that another mechanism, which they term generically the bias of technical change, must be at work.

Lawrence (2015) is also concerned with studying the fall in the US labor share through the capital-labor substitution channel; his study presents evidence that growth in human capital, and hence in the quantity of effective hours supplied by labor, has been so large as to *reduce* capital-labor ratios, which has the effect of reducing the labor share when $\sigma < 1$.

The rest of the paper is organized as follows. Section 2 develops a static model that presents, in the simplest possible setting, the main analytical results in the paper: the connection between moments of the size distribution of firms operating capital with differing degrees of automation to the parameters of an aggregate CES production function. Section 3 embeds this static production structure into a dynamic model which, in aggregate, closely resembles the standard neoclassical growth model, and derives the requirements for existence and properties of a balanced growth path of the economy. In Section 4 I present the empirical approach; Section 5 presents evidence of the model's fit to US industry-level data and the implications for the impact of automation on the US industry-level and aggregate labor shares in the past 40 years. Section 6 concludes.

2 Static Model

I start with a static version of the model in order to simplify the exposition of the firm level technology and the aggregation result that yields a constant elasticity of substitution (CES) macro production function. The static economy consists of a continuum of firms, each operating a firm-specific technology to produce a homogeneous good that is sold in a competitive market. To produce, firms must hire labor, which is supplied by homogeneous workers in a competitive labor market.

2.1 Firms

A continuum of competitive firms produce a unique final good Y. A firm i is characterized by its technology $\tilde{a}_i$ and capital stock k_i. Technology is embodied in the capital stock, so a firm's capital consists of k_i units of capital that embody the technology level $\tilde{a}_i$. I assume that each firm owns exclusive rights to its technology, so there is one firm per technology $\tilde{a}_i$.

Technology Production of one final good requires the completion of a continuum of all tasks in an interval $[0, q]$[3]. Tasks can be produced by both workers and machines. Workers can perform all tasks in $[0, q]$,[4] whereas a machine of type $\tilde{a}$ can perform tasks in $[0, \tilde{a}]$, with

[3]Following Becker and Murphy (1992), the production function *in terms of tasks* is the Leontieff function,

$$y = \min_{0<x<q} y(x)$$

with the interpretation that producing y goods requires that each task x be performed $y(x)$ times.

[4]Assuming that workers can perform all tasks is not essential to the results but simplifies notation and algebra. What is important is that there are some tasks that only workers can perform.

$\tilde{a} < q$. So to produce one unit of the final good using a machine of type $\tilde{a}$, a worker has to complete the remaining tasks $[\tilde{a}, q]$. If it takes workers one unit of time to complete all tasks in $[0, q]$, a worker-machine pair would produce $1/(q - \tilde{a})$ goods in one unit of time; I further assume that the productivity of a worker-machine pair is increasing in the quantity of tasks, and that the productivity gains from worker specialization are governed by a parameter $\gamma > 1$. With these assumptions, the productivity of a worker-machine pair is given by:

$$z(a) = \left(\frac{q}{q - \tilde{a}}\right)^{\gamma} = \left(\frac{1}{1 - a}\right)^{\gamma} \tag{1}$$

Where I define $a \equiv \frac{\tilde{a}}{q}$ as the fraction of total tasks that the machine can complete, which I call the machine's (and firm's) *degree of automation.* When referring to machine types, I will adopt the notational convention that variables with a tilde measure the highest *level* of task that a machine can perform, whereas variables without a tilde measure *the fraction of total production tasks* that a machine can perform (this distinction will become critical in the dynamic version of the model). The function $z(a)$ is related to the span of control in models in which production is organized in knowledge hierarchies as in Garicano (2000) and Garicano and Rossi-Hansberg (2006). In those papers, productivity gains arise from employing more knowledgeable workers at higher levels in hierarchy of production. Here, productivity gains come from pairing workers with machines because workers are essential to production (since $\tilde{a} < q$) and machines save worker time by performing a portion of the tasks required in production. The degree of productivity gains are scaled by the parameter γ, which I interpret in the spirit of Becker and Murphy (1992) as capturing returns to worker specialization.

Firm production function and profit maximization Each firm i is characterized by its technology $\tilde{a}_i$ (equivalently, its degree of automation a_i) and capital stock k_i. At the firm level, I assume that the number of (effective) worker hours per machine is technologically constrained to be in fixed proportion, resulting in the following (Leontieff) production function for firm i:

$$y_i = z(a_i) \min(k_i, H \cdot n_i), \tag{2}$$

where k_i and n_i are the firm's capital stock and labor input. H is labor-augmenting productivity, common to all workers, so that $H \cdot n_i$ is the efficiency adjusted number of hours worked by firm i's workers.

Firm size distribution I assume that each firm has exclusive use of a technology, so there is a one-to-one correspondence between i and a_i, and I can model the firm size distribution as the measure of capital at each degree of automation a_i. I denote this measure $k(a)$, and assume that it has bounded support $a \in [a_\ell, a_h]$, so that a_ℓ and a_h are, respectively, the least and most automated types of capital in the economy. Since a is a fraction and I have assumed that at least some level of worker-only tasks are needed in production, the bounds satisfy $0 \leq a_\ell < a_h < 1$. Anticipating the discussion of aggregation in the following section, the integral of $k(a)$ is the aggregate capital stock of the economy, $K = \int_{a_\ell}^{a_h} k(a)\ \mathrm{d}a$. Normalizing the firm size distribution by aggregate capital, I define the PDF $f(a) \equiv \frac{k(a)}{K}$, the density function of degrees of automation in the economy. Further, I define the density $g(z)$ as the distribution of the function $z(a)$ given that a has density $f(a)$. The density function $g(z)$ is the measure of productivities associated with a given distribution of automation. Since $f(a)$ is bounded, so too is g(z), with lower bound $z_\ell \equiv z(a_\ell)$ and upper bound $z_h \equiv z(a_h)$.

2.2 Aggregation and Equilibrium

I start by solving for aggregate labor demand. The profit maximization problem for a firm with productivity z is:

$$\max_n \ z \min\left(k(z), H \cdot n\right) - w \cdot n \tag{3}$$

It follows that the optimal choice of hours as a function of z is :

$$n(z) = \begin{cases} \frac{k(z)}{H} & \text{if } z \geq \frac{w}{H} \\ 0 & \text{if } z < \frac{w}{H} \end{cases}$$

Since z is a function of a, I can rewrite the optimal choice of hours in terms of the degree of automation of the marginal firm.

$$n(a) = \begin{cases} \frac{k(a)}{H} & \text{if } a \geq a_w \\ 0 & \text{if } a < a_w \end{cases}$$

Where $a_w = 1 - \left(\frac{w}{H}\right)^{\sigma-1}$.

Aggregate labor demand, N, is the integral over firm labor choices, $N = \int_{z_\ell}^{z_h} n(z)\ \mathrm{d}z$. Substituting in the optimal choice of hours, and $k(z) = K \cdot g(z)$, the integral becomes $N = \frac{K}{H}\int_{w/H}^{z_h} g(z)\ \mathrm{d}z$. Denoting by $G(\cdot)$ the CDF of z, integrating gives aggregate labor demand as:

$$N = \frac{K}{H}\left[1 - G\left(\frac{w}{H}\right)\right] \tag{4}$$

Aggregate labor demand is the product of the total capital stock, divided by worker efficiency, times the fraction of active firms (1 minus the CDF at $\frac{w}{H}$, the term in square brackets). I assume that workers' preferences give a labor supply function $N^s(w)$, increasing in the wage; equating $N^s(w)$ to labor demand as in Equation 4 gives the equilibrium wage. With the equilibrium wage in hand, aggregate output Y can be solved for as the integral over firm-level output, $Y = \int_{w/H}^{z^h} zk(z)\ \mathrm{d}z = K\int_{w/H}^{z^h} zg(z)\ \mathrm{d}z$, which can be expressed as:

$$Y = K\left[1 - G\left(\frac{w}{H}\right)\right]\mathbb{E}_G\left(z|z \geq \frac{w}{H}\right) \tag{5}$$

Aggregate output is the product of three terms: the capital stock, the fraction of active firms, and the average productivity of active firms, given by the third term ($\mathbb{E}_G\left(z|z \geq \frac{w}{H}\right) = \left[1 - G\left(\frac{w}{H}\right)\right]^{-1}\int_{w/H}^{z^h} zg(z)\ \mathrm{d}z$).

2.3 From Automation to a CES Aggregate Production Function

Equation 5 gives aggregate output as a function of the equilibrium wage. In order to express aggregate output in a more familiar form, as a function of capital, labor and technology, I start with an assumption about the distribution of degrees of automation $f(a)$. In particular, I assume that $f(a)$ is a 3 parameter beta distribution[5], characterized by one shape parameter, ρ, and lower and upper bounds a_ℓ and a_h. The following lemma characterizes the distribution of $z(a)$ using this assumption.

Lemma 1. *If $a \sim B(1, \rho; a_\ell, a_h)$, so that the PDF of a, $f(a) = \rho\frac{(a_h - a)^{\rho-1}}{(a_h - a_\ell)^\rho}$, then $z(a) =$*

[5] The standard Beta distribution is bounded in $[0, 1]$ and characterized by two shape parameters. I restrict one of the shape parameters to equal 1 and set the bounds to $0 < a_\ell < a_h < 1$.

$\left(\frac{1}{1-a}\right)^{\gamma}$ *is distributed with CDF* $G(z)$*:*

$$G(z) = 1 - \left(1 - \frac{1 - \left(\frac{z_\ell}{z}\right)^{\frac{1}{\gamma}}}{1 - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}}\right)^{\rho} \tag{6}$$

and $z \in [z_\ell, z_h]$*, with* $z_\ell = \left(\frac{1}{1-a_\ell}\right)^{\gamma}$ *and* $z_h = \left(\frac{1}{1-a_h}\right)^{\gamma}$*.*

Proof. See Appendix A.1 □

With $\gamma = 1$, this lemma is equivalent to the derivation in Geerolf (2017) of a Pareto-distributed span of control in an economy with a simplified Garicano (2000) production function and a Beta distribution for worker skills. The distribution $G(z)$, known to statisticians as a truncated Beta-Pareto distribution[6], is closely related to the Pareto distribution; indeed with $\rho = 1$, $G(z)$ is exactly a Pareto distribution with tail parameter $\frac{1}{\gamma}$ and lower bound z_ℓ, truncated at an upper bound z_h.

The distribution $G(z)$ generalizes the distribution derived by Sato (1969) (which in turn extends the result in Levhari (1968)) for $z_\ell > 0$. In contrast to those papers, I do not start with $G(z)$ as a primitive, but instead derive it from the underling distribution $F(a)$ and productivity function $z(a)$. These differences are crucial since they allow me to develop a theoretical link between automation and the aggregate production function, and are also important in proving the existence of a balanced growth path and its associated stationary distribution in the dynamic model presented in Section 3.

Before proceeding I make a further restriction on parameters: for the remainder of the paper, I assume that $\gamma = \rho + 1$. This is not an innocuous restriction, since in the logic of the model γ can be thought of as a deep technological parameter that governs firm-level returns to automation, whereas ρ is the shape parameter of the automation distribution, which makes it closer to an equilibrium object. The gain from imposing this restriction is substantial: as shown in Proposition 1, with this parametrization aggregate output in the economy can be represented exactly as a canonical CES production function. The connection between a knife edge parametrization of the model and the aggregation result may seem mysterious, but is in fact rather prosaic: it is only in the knife-edge case that the integral in Equation

[6] Akinsete et al. (2008) and Lorenzutti et al. (2012) derive moments, MGF, etc for this distribution. I show some of its properties in Appendix C.

5 can be solved in closed form. I have verified numerically that the results presented in this section and the next are not substantially affected by deviations from the knife edge parametrization. To economize on notation and consistency with the literature, I also define $\rho \equiv \frac{\sigma}{1-\sigma}$ (σ will correspond exactly to the elasticity parameter of the CES). In sum, I make the following assumptions about the distributions $F(a)$ and $G(z)$:

Assumption 1: Distributions $F(a)$ and $G(z)$ *The automation distribution $F(a)$ and its associated productivity distribution $G(z)$, with $z(a) = \left(\frac{1}{1-a}\right)^{\frac{1}{1-\sigma}}$, are:*

$$F(a) = \left(\frac{a_h - a}{a_h - a_\ell}\right)^{\frac{\sigma}{1-\sigma}} \qquad \text{and} \qquad G(z) = 1 - \left(1 - \frac{1 - \left(\frac{z_\ell}{z}\right)^{1-\sigma}}{1 - \left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{1-\sigma}}$$

The following proposition shows that, given this assumption, aggregate output can be represented by a CES production function with elasticity of substitution $0 < \sigma < 1$.

Proposition 1. *Using Assumption 1, aggregate output for this economy can be represented as the CES production function with elasticity of substitution $0 < \sigma < 1$,*

$$Y = A\left(\alpha K^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(H \cdot N)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{7}$$

where H is labor-augmenting productivity and the capital (K) and labor aggregates (N) are:

$$K = \int_Z k(z)\ dz \qquad \text{and} \qquad N = \frac{K}{H}\left[1 - G\left(\frac{w}{H}\right)\right]$$

where w is the market clearing wage that solves $N^s(w) = N$. The total factor productivity term A and capital distribution parameter α are functions of parameters of the productivity distribution $G(z)$:

$$A = (z_h)^{1-\sigma}(z_\ell)^{\sigma} \tag{8}$$

$$\alpha = 1 - \left(\frac{z_\ell}{z_h}\right)^{1-\sigma} \tag{9}$$

Proof. See Appendix A.2. □

Proposition 1 contains 3 linkages between the parameters of the automation distribution $F(a)$ and the aggregate production function. I discuss each of these in turn.

Elasticity of substitution σ The elasticity of substitution σ of the aggregate production function is related to the shape parameter of the underlying beta distribution[7]. Inspecting the CDF of a $B\left(1, \frac{\sigma}{1-\sigma}\right)$ distribution, $F(x) = (1-x)^{\frac{\sigma}{1-\sigma}}$, the effect of increasing σ is to shift mass towards the left tail of the distribution, which in the model means that firms with relatively low automation are relatively larger. The intuition for why this translates into a higher elasticity of substitution is that the higher is σ, the larger is the mass of firms in the neighborhood of the cutoff level of automation a_w, and consequently the higher is the mass of firms that go from active to idle in response to a small increase in the cutoff a_w (corresponding to a small increase in the wage). Because firms in the neighborhood of this cutoff require more labor per unit of output, in the aggregate the economy becomes more capital intensive when these firms shut down. In other words, the aggregate economy substitutes towards capital in response to an increase in the wage, and this effect is stronger the higher is σ.

The aggregation result in Proposition 1 requires that σ be smaller than 1. The discussion in the previous paragraph provides a hint as to why that is the case: $\sigma > 1$ would imply that the economy becomes *more* labor intensive as the wage increases. Sato (1969) shows that this is a general existence requirement in this class of aggregative models (Mangin (2015) proves a similar result using a very different approach). This is an interesting theoretical insight in the long-running debate about the value of σ in a macro model with an aggregate production function.

A and α The TFP parameter A and distribution parameter α are functions of moments of the productivity distribution $G(z)$; as shown in the appendix, A is the mean of z and as such is increasing in both the lower and upper bounds of the productivity distribution. TFP is therefore increasing in both a_h and a_ℓ: the more automated the capital stock, the higher is total factor productivity. Substituting for z_h and z_ℓ, the distribution parameter α

[7] I have imposed $\gamma = \frac{1}{1-\sigma}$ so strictly speaking σ is also related to the returns to automation at the firm level. In describing the effect of varying σ I have in mind a local approximation leaving γ unchanged and varying the shape parameter of the Beta distribution. The resulting aggregate production function is not an exact CES function, but is very close to CES for small deviations.

can be expressed in terms of a_h and a_ℓ, $\alpha = \frac{a_h - a_\ell}{1 - a_\ell}$. The distribution parameter is therefore increasing in a_h and decreasing in a_ℓ - I relegate the interpretation of this result to the discussion of Proposition 2 below.

Proposition 2. *The labor share in this economy is given by*

$$LS = \frac{1 - a_h}{1 - a_w} \tag{10}$$

Proof. The labor share with a CES production function can be expressed as:

$$LS = (1 - \alpha)^\sigma A^{\sigma - 1} \left(\frac{w}{H}\right)^{1-\sigma}$$

Substituting in for A and α from Equations 8 and 9, and replacing $z_h = \left(\frac{1}{1-a_h}\right)^{\frac{1}{1-\sigma}}$ and $\frac{w}{H} = \left(\frac{1}{1-a_w}\right)^{\frac{1}{1-\sigma}}$ gives the result after some algebraic manipulation. □

Proposition 2 provides an easily interpretable expression for the labor share in terms of the task distribution: the labor share is given by the ratio of the fraction of worker tasks at the most automated firm $(1 - a_h)$ over the fraction of worker tasks at the least automated firm $(1 - a_w)$.

Both the numerator and denominator of this ratio can be thought of as measures of automation, but capturing different margins. An increase in a_h can be interpreted as an *extensive* margin of automation, since an increase in a_h implies that new automation technologies have been discovered and embodied as capital goods. On the other hand, an increase in a_w can be interpreted as an increase along an *intensive* margin: some fraction of the least automated technologies installed in the economy is left idle, which means that the average automation level of firms that are still active increases.

This interpretation is related to the dynamics at work in models of innovation and diffusion. Increasing a_h is analogous to innovation - a jump in the technology frontier. Increasing a_w corresponds to diffusion, since the technology of the lowest active firm is closer to the frontier after an increase in a_w. To make the analogy to the diffusion literature more specific: Benhabib et al. (2017) develop a model in which the firm productivity distribution has a finite upper bound, so that both innovation (growth at the frontier) and diffusion (catch up from firms in the left tail of the distribution) are present; they find that the presence of a finite upper bound affects the key properties of the BGP equilibrium. By analogy, the model in this paper shares uses the same firm level technology as Lagos (2006), but in the models of that paper firm productivities are not bounded above (so only the "diffusion" margin is active). At a conceptual level, adding the upper bound allows me to derive model of both the TFP term A *and* capital share parameter α of the aggregate production function.

3 Dynamic Model and Balanced Growth

In this section, I embed the static structure described in the previous section into a dynamic model that, in the aggregate, closely resembles a standard neoclassical growth model. I add households to the model; as usual, households supply labor and make consumption and savings decisions that determine the evolution of the capital stock.

I introduce growth into the model by assuming exogenous growth paths for three variables: the number of tasks required for the production of the final good, q_t; the frontier level of automation technology $\tilde{a}_{h,t}$ and labor productivity H_t. In addition to these three exogenous states, the endogenous aggregate states of the dynamic economy are the aggregate capital stock K_t and the distribution of automation technology $F_t(a)$.

The set up of the dynamic model is such that the distribution $F_t(a)$ stays within the Beta family that allows aggregation as per Proposition 1, which reduces the state vector from the entire distribution $F_t(a)$ to the three parameters that characterize the Beta: the shape parameter σ and the upper $\left(a_{h,t} = \frac{\tilde{a}_{h,t}}{q_t}\right)$ and lower $\left(a_{\ell,t} = \frac{\tilde{a}_{\ell,t}}{q_t}\right)$ bounds of the automation distribution. Further, the assumptions will also keep the shape parameter σ constant, so that the distribution can be summarized by its upper and lower bounds. From Proposition 1, this will ensure that the aggregate production function will be CES with elasticity of substitution σ, but the TFP parameter A_t and capital distribution parameter α_t will be state dependent through their dependence on the bounds of the automation distribution.

The production and investment technology in the dynamic model are closely related to models with putty-clay technology and irreversible investment, introduced by Johansen (1959). The vintage capital structure in a putty-clay model with irreversible investment is a natural way to model an economy in which different types of capital coexist in the economy at any given time. In this model, a type or vintage of capital corresponds to capital with different levels of automation. The implementation of putty-clay technology in this model is closely related to that of Gilchrist and Williams (2000): the firm-level Leontieff production function leads to variable utilization, and old (less automated) capital is scrapped endogenously as newer (more automated) capital in installed.

3.1 Firms: entry, exit and investment

Firms are characterized by two state variables: a technology level $\tilde{a}_i$ and a firm-specific capital stock $k_{i,t}$ (for simplicity I refer to a unit of capital as a machine). I start by describing the production and investment decisions of an incumbent firm, followed by the endogenous exit decision, and finally entry of new firms.

Incumbent firms A firm is characterized by its level of automation $a_{i,t} = \frac{\tilde{a}_i}{q_t}$ and capital stock $k_{i,t}$ and produces final output by combining capital and labor $(n_{i,t})$ in the Leontieff production function $y_{i,t} = z(a_{i,t}) \min[k_{i,t}, H_t \cdot n_{i,t}]$, where $z(a_{i,t}) = \left(\frac{1}{1-a_{i,t}}\right)^{\frac{1}{1-\sigma}}$. Given the Leontieff technology, the firms' optimal choice of hours is given by a cutoff rule (as in the static model of Section 2.2): firms with productivity $z(a_{i,t}) < \frac{w_t}{H_t}$ do not hire labor in period t, so their capital remains idle. The firm's profit per unit of capital is therefore given by:

$$\pi(a_{i,t}) = \max\left[z(a_{i,t}) - \frac{w_t}{H_t}, 0\right] \tag{11}$$

A firm's technology level $\tilde{a}_i$ is fixed when it enters (firm entry is described in detail below). Each firm has a monopoly on the creation of new machines that embody that technology. However, because the number of tasks required to produce final output, q_t, changes over time, a firm's degree of automation, $a_{i,t} = \frac{\tilde{a}_i}{q_t}$ is not constant over time. In particular, the degree of automation of a machine that can perform tasks up to $\tilde{a}_i$ decreases over time if q_t grows, since that machine is less automated relative to the number of tasks that workers perform. This gives rise to technological obsolescence and endogenous scrapping (described below). The law of motion for $a_{i,t}$ is:

$$a_{i,t+1} = \frac{\tilde{a}_i}{q_{t+1}} = \frac{q_t}{q_{t+1}} a_{i,t} \tag{12}$$

At the end of each period, the firm has to pay a maintenance cost $\kappa > 0$ per machine to keep a machine operational. If the firm does not pay the maintenance cost, the machine depreciates completely at the end of the period; in addition, machines fail with i.i.d. probability δ. Firms can increase their capital stock by creating new machines of their firm-specific type. To do so they face a quadratic cost function. The value function of a firm with capital $k_{i,t}$ and automation $a_{i,t}$, $V(a_{i,t}, k_{i,t})$ that chooses to pay the maintenance cost κ is given by:

$$V(a_{i,t}, k_{i,t}) = \max_{\iota_t} \ \pi(a_{i,t}) k_{i,t} - \kappa k_{i,t} - \frac{\phi(a_{i,t})}{2} k_{i,t} \left(\frac{\iota_{i,t}}{k_{i,t}}\right)^2 + R_{t+1} V(a_{i,t+1}, k_{i,t+1})$$

$$s.t.$$

$$\iota_t \geq 0$$

$$k_{i,t+1} = (1-\delta) k_{i,t} + \iota_{i,t}$$

Where ι_t is the firms' investment, R_{t+1} is the household's discount factor, and $\phi(a_{i,t})$ is an increasing, convex function such that the marginal cost of creating machines is increasing in the level of automation. The function ϕ determines the quantity of investment at each level of automation, so is important in determining the evolution of the distribution of automation. I return to this point in Section 3.2.

The continuation value of the machine is the discounted value net of the maintenance cost. Incumbent firms with machines that have positive (net) continuation value invest and grow their capital stock. The value function of a firm with $k_{i,t}$ type $a_{i,t}$ machines is:

$$\max_{\iota_{i,t}} \ R_{t+1} V(a_{i,t+1}, k_{i,t+1}) - \frac{\phi(a_{i,t+1})}{2} k_{i,t} \left(\frac{\iota_{i,t}}{k_{i,t}}\right)^2 \tag{13}$$

$$\text{s.t. } \iota_{i,t} \geq 0$$
$$k_{i,t+1} = (1-\delta) k_{i,t} + \iota_{i,t}$$

With this specification for the investment adjustment cost, the value function for a firm is homogeneous of degree one in the firm capital stock, so that the value of the firm is simply the value of each of its machines multiplied by its capital stock. $V(a,k) = k \cdot v(a)$ The FOC for the investment decision is:

$$\phi\left(a_{i,t+1}\right)\iota_{i,t} = R_{t+1}v\left(a_{i,t+t}\right) \tag{14}$$

which gives that at the optimum level of investment, the marginal cost of creating a new machine is equated to the marginal benefit, which is the value of the new machine. Since the discounted value on the right hand side is strictly positive (otherwise machines of that type would have been scrapped), and $0 < \phi\left(a_{t+1}\right) < \infty$, $\iota_{i,t} > 0$. This implies that the non-negativity constraint does not bind, and that any firm that does not scrap its capital has strictly positive investment. The law of motion for firm-specific capital is:

$$k_{i,t+1} = (1-\delta)\,k_{i,t} + \iota_{i,t} \tag{15}$$

Scrapping and firm exit The scrapping threshold at the end of the period is defined implicitly by

$$(1-\delta)\,R_{t+1}v\left(a_{\ell,t+1}\right) = \kappa \tag{16}$$

The automation level of the lowest machine installed at the beginning of time t is $a_{\ell,t}$, so the scrapping threshold condition implies that machines in $(a_{\ell,t}, a_{\ell,t+1}]$ are scrapped each period, and the firms that own those technologies exit. In terms of the CDF of the automation distribution, the fraction of capital that is scrapped at the end of time t is $\varsigma_t = F_t\left(a_{\ell,t+1}\right)$.

New automation technologies and firm entry At the end of each period, new automation technologies $(\tilde{a}_{h,t}, \tilde{a}_{h,t+1})$ are discovered. A large mass of potential entrants each draw a technology from this set, such that $\tilde{a}_{h,t+1} - \tilde{a}_{h,t}$ new firms are created each period, each with a unique $\tilde{a}_i$. These firms enter with zero capital and solve the following initial investment decision:

$$\max_{k_{i,t+1}}\; k_{i,t+1}R_{t+1}v\left(a_{i,t+1}\right) - \frac{\phi\left(a_{i,t+1}\right)}{2}\left(k_{i,t+1}\right)^2 \tag{17}$$

The first order condition is given by:

$$\phi\left(a_{i,t+1}\right)k_{i,t+1} = \beta R_{t+1}v\left(a_{i,t+1}\right) \tag{18}$$

3.2 Dynamics of the automation distribution $F_t(a)$

The evolution of the automation distribution $F_t(a)$ is determined by the investment decisions of incumbents and new entrants, and the scrapping decision of exiting firms. Denoting by $J_t(a)$ the distribution of investment, the law of motion of $F(a)$ in general is a complicated map from $(F_t(a), J_t(a))$ to $F_{t+1}(a)$. I vastly reduce this complexity by making assumptions that ensure that the aggregate production function of the economy is a CES function with invariant elasticity of substitution at all times, which in turn implies that $F_t(a)$ is a beta distribution with invariant shape parameter for all t. Further, given the assumption of exogenous growth in $\tilde{a}_{h,t}$ and q_t, the upper bound of F_t evolves exogenously. The problem of tracking the automation distribution is therefore reduced to the law of motion of the lower bound $a_{\ell,t}$ of the distribution.

The necessary assumptions are implemented through the functions $\phi(a)$ and $\phi^N(a)$, the marginal cost of capital for incumbents and new entrants, respectively. Since I will not be studying the dynamics of the model outside the balanced growth path, I do not specify functional forms for $\phi(a)$ and $\phi^N(a)$ outside the steady state.

3.3 Households

The representative household chooses consumption C_t, investment I_t and labor supply N_t every period to maximize the following intertemporal problem:

$$\max_{C_t, I_t, N_t} \sum_{t=0}^{\infty} \beta^t U(C_t, N_t) \tag{19}$$

Subject to the budget constraint:

$$C_t + P_{I,t} I_t = w_t N_t + r_{K,t} K_t \tag{20}$$

Where $P_{I,t}$ is the price of investment goods and $r_{K,t}$ is the rental rate of capital. The equilibrium conditions from household optimization are the standard intratemporal labor supply condition and an intertemporal Euler equation:

$$U_{N,t} = -U_{C,t} w_t \tag{21}$$

$$U_{C,t} = \beta \left[U_{C,t+1} \left(\frac{P_{I,t+1} (1-\delta) (1-\varsigma_{t+1}) + r_{K,t+1}}{P_{I,t}} \right) \right] \tag{22}$$

Where the utility function $U(C, N)$ satisfies the conditions required for balanced growth (as in King et al. (1988)). The aggregate capital stock has the following law of motion:

$$K_{t+1} = (1-\delta)(1-\varsigma_t) K_t + I_t \tag{23}$$

Where δ is exogenous depreciation and ς_t is the fraction of the capital stock that is scrapped every period due to obsolescence.

3.4 Equilibrium and balanced growth

An equilibrium of the dynamic model consists of sequences of: (i) prices $\{w_t,\ P_{I,t},\ r_{K,t}\}$; (ii) aggregate allocations $\{C_t,\ I_t,\ N_t\}$; (iii) firm allocations $\{\iota_{i,t},\ n_{,it}\}$; and (iv) firm exit decisions, such that household and firm equilibrium conditions are satisfied and markets for the final good and savings clear. Market clearing in final goods requires that total final goods output is either consumed or invested:

$$Y_t = C_t + P_{I,t} I_T \tag{24}$$

and savings market clearing requires that savings $P_{I,t} I_t$ equal aggregate investment expenditures:

$$P_{I,t} I_t = \int_{a_{\ell,t+1}}^{a_{h,t+1}} p_t(a)\, \iota_t(a)\ \mathrm{d}a \tag{25}$$

Where $\iota_t(a)$ is the investment choice of a firm with automation level a.

Balanced Growth I now show that the model has a steady state with constant growth that satisfies the Jones and Scrimgeour (2008) definition of a balanced growth path. Namely, quantities $\{Y_t, C_t, I_t, K_t\}$ grow at constant rates and factor shares are constant and strictly positive. I add to the usual definition of the balanced growth path the requirement that the automation distribution $F(a)$ is stationary. In fact, as shown below, stationarity of $F(a)$ is a necessary and sufficient condition for a balanced growth path.

Since aggregate output in this economy can be represented as the following CES production function with elasticity of substitution $\sigma < 1$:

$$Y_t = A_t \left(\alpha_t K_t^{\frac{\sigma-1}{\sigma}} + (1-\alpha_t)(H_t \cdot N_t)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

the conditions for the existence of a balanced growth path follow from the Uzawa (1961) growth theorem, which states that the balanced growth path exists only if all technological growth is labor augmenting. This requires that the TFP term $A_t = (z_{h,t})^{1-\sigma} (z_{\ell,t})^{\sigma}$ is constant in steady state. Substituting in the definitions for z_h and z_ℓ in terms of the bounds of the automation distribution $F(a)$ gives $A_t = \left(\frac{1}{1-a_{h,t}}\right)\left(\frac{1}{1-a_{\ell,t}}\right)^{\frac{\sigma}{1-\sigma}}$. Constant TFP therefore requires that $a_h = \frac{\tilde{a}_{h,t}}{q_t}$ and $a_\ell = \frac{\tilde{a}_{\ell,t}}{q_t}$ are both constant in the steady state, so q_t, $\tilde{a}_{h,t}$, $\tilde{a}_{\ell,t}$ must grow at the same rate; and that σ is constant. Since a_h, a_ℓ and σ fully characterize the distribution $F(a)$, the conditions for constant TFP in the steady state and stationarity of the distribution $F(a)$ are the same.

Assuming that the labor-augmenting factor H_t grows at a constant growth rate $1+g = \frac{H_{t+1}}{H_t}$, in the balanced growth path $\{Y_t, C_t, I_t, K_t\}$ also grow at rate g. This implies that there is no long run growth from automation, and that the distribution of automation $F(a)$ does not affect the steady state growth rate. The automation distribution does, however, determine factor shares in the balanced growth path. From Proposition 2, the labor share is $LS = \frac{1-a_h}{1-a_w}$. Since balanced growth is consistent with *any* a_h and a_w (as long as $0 < a_w < a_h < 1$), the model has a continuum of balanced growth paths indexed by the steady state labor share. In terms of the parameters of CES aggregate production function, there is a balanced growth path for any $\alpha \in (0,1)$.

At a conceptual level the model can therefore rationalize trends in factor shares as transitions between balanced growth paths brought about by periods of unbalanced growth in the number of tasks required in production, q_t, and the frontier automation technology $\tilde{a}_{h,t}$.

4 Measuring Automation in the Data

In this section I show how to take the model of automation presented above to data, with the objective of measuring the contribution of automation to changes in labor's share of US GDP. To do so I apply the model described in Section 3 to industry-level data. I will think of each industry in the US economy as consisting of firms operating firm-specific technologies

Decade	ΔLS_{US}	Shift	Share
1972-80	-0.01	-0.00	-0.01
1980-90	0.02	0.01	0.01
1990-00	0.01	0.00	0.01
2000-10	-0.06	-0.06	0.00

Source: US KLEMS. Calculations exclude public sector and FIRE.

Table 1: Shift and Share Components of Change in US Labor Share, 1972-2010

to produce an industry-specific good, and of the wage in each industry as being determined in an industry-specific labor market. In the background, preferences and the investment technology aggregate the output of the different industries into a final consumption and investment good, but I do not model these explicitly; instead, to aggregate the output and labor shares of all industries I use output shares directly from the data. Denoting by YS_j an industry's share of overall output and by LS_j an industry's labor share, one can decompose the change in the aggregate labor share into "shift" and "share" components around expansion points $\overline{YS}_j$ and $\overline{LS}_j$:

$$\Delta LS_{US} = \underbrace{\sum_j \overline{YS}_j \Delta LS_j}_{Shift} + \underbrace{\sum_j \overline{LS}_j \Delta YS_j}_{Share} \tag{26}$$

The shift component measures the contribution of within-sector changes, and the share component measures the contribution of cross-industry reallocation. The objective of the empirical exercise in Section 5 is to explore the model's ability to explain the shift component, taking the share component from the data.

As Table 1 shows, the share component is certainly an important one: trends in the industrial composition of the US economy have a first order effect on the macro aggregates that I study. But explaining these trends is a significant challenge in its own right and beyond the scope of this paper.

Applying Equation 10 to each sector j, the labor share expressed in terms of the frontier (a_h) and marginal (a_w) automation technologies is:

$$LS_{j,t} \equiv \frac{w_{j,t} N_{j,t}}{Y_{j,t}} = \frac{1 - a_{h,j,t}}{1 - a_{w,j,t}} \tag{27}$$

I detail below the steps taken to take this sharp prediction implied by the model to the data.

To do so, I combine macro- and micro-data at the industry level with detailed occupation-level task data to measure task inputs into production at the industry level. I start by deriving model measurements of total and machine task inputs into production before describing the measurement of these objects in the data.

4.1 Task production in the model

I start by deriving model expressions for the aggregate amount of worker and machine tasks required to produce a given level output.

Starting at the firm level, to produce one unit of output, workers at a a firm operating capital that can complete tasks up to $\tilde{a}_i$ have to complete tasks $q_t - \tilde{a}_i$[8]. Firm i's workers therefore complete a total of

$$\lambda_{i,t} \equiv y_{i,t} \left(q_t - \tilde{a}_i \right) = y_{i,t} q_t \left(1 - a_{i,t} \right) \tag{28}$$

tasks. Firm i's machines complete the complementary fraction of tasks,

$$\mu_{i,t} \equiv y_{i,t} q_t a_{i,t} \tag{29}$$

Total task production at the firm level $\lambda_{i,t} + \mu_{i,t} = q_t y_{i,t}$ is increasing in the total number of tasks required for production and in the quantity of output. To obtain aggregate task inputs, I integrate $\lambda_{i,t}$ and $\mu_{i,t}$ over the set of firms. In Appendix B.1, I derive the following expressions for total worker (Λ_t) and capital (M_t) task inputs per unit of effective labor input:

$$\Lambda_t = q_t \mathbb{E}_{F,t} \left[1 - a | a > a_{w,t} \right] \tag{30}$$

$$M_t = q_t \mathbb{E}_{F,t} \left[a | a > a_{w,t} \right] \tag{31}$$

The expressions are intuitive: the total number of worker (capital) tasks per effective worker hour ($H_t \cdot N_t$) is the total number of tasks q_t times the average fraction of worker (capital) tasks among active firms (firms with $a > a_{w,t}$).

[8] To economize on subscripts, I omit the industry subscript j except when necessary to avoid confusion.

4.2 Model measurement equations

Assuming we can observe Λ_t and M_t in the data, Equations 30 and 31 form a system of two observables and three latent variables: q_t, $a_{h,t}$ and $a_{w,t}$. Taking the ratio $\frac{\Lambda_t}{M_t}$ eliminates one of these latest variables, q_t, since the ratio of worker to machine tasks doesn't scale with the total number of tasks q_t. Taking this ratio as a departure, I show in Appendix B.1 how to derive the following measurement equation by making a log-linear approximation around $a_{h,t-1}$ and $a_{w,t-1}$

$$\Delta \log \Lambda_t - \Delta \log M_t = \chi_{h,t} \Delta \log a_{h,t} + \chi_{w,t} \Delta \log a_{w,t} \tag{32}$$

Where Δ denotes the time difference operator and $\chi_{h,t}$ and $\chi_{w,t}$ are time dependent coefficients (the time dependence comes from successively approximating around the lagged values $a_{h,t-1}$ and $a_{w,t-1}$). Equation 32 is one measurement equation that I will take to the data. To obtain a second, I start with the definition of $a_{w,t}$ as the level of automation of the marginal active firm, $\frac{w_t}{H_t} = \left(\frac{1}{1-a_{w,t}}\right)^{\frac{1}{1-\sigma}}$. Solving for $1 - a_{w,t}$ gives:

$$1 - a_{w,t} = \left(\frac{w_t}{H_t}\right)^{\sigma - 1} \tag{33}$$

Assuming $\Delta \log \Lambda_t$, $\Delta \log M_t$, w_t and H_t are observable, Equations 32 and 33 form a system of two measurement equations and two latent variables, $a_{w,t}$ and $a_{h,t}$. The following section describes the measurement of the required variables in the data.

4.3 Data

I use three data sources to measure wages ($w_{j,t}$), labor augmenting productivity ($H_{j,t}$), workers tasks ($\Delta \log \Lambda_{j,t}$), and machine tasks ($\Delta \log M_{j,t}$). Table 2 summarizes the data used in the empirical exercise.

Industry wages $w_{j,t}$ and labor input quality $H_{j,t}$ come from the World KLEMS database for the US, as described in Jorgenson et al. (2012). I also use KLEMS data to construct industry-level labor and value added shares. I measure $\Delta \log \hat{a}_{w,j,t}$ by taking a first order

	Data	Source
$\hat{w}_{j,t}$	Real hourly wage	KLEMS
$\widehat{H}_{j,t}$	Labor services per hour	KLEMS
$\widehat{\Lambda}_{j,t}$	Total task inputs	CPS and O*NET
$\widehat{\Lambda}_{M,t}$	Automatable task inputs	CPS and O*NET

Table 2: Data Used in Empirical Exercise

approximation of Equation 33, and using KLEMS data on wages and labor services per hour.

$$\Delta \log \hat{a}_{w,j,t} = (1-\sigma)\,\frac{1-\hat{a}_{w,j,t-1}}{\hat{a}_{w,j,t-1}}\left(\Delta \log \hat{w}_{j,t} - \Delta \log \widehat{H}_{j,t}\right) \tag{34}$$

4.3.1 Measuring task inputs into production

To measure task inputs into production, I combine Current Population Survey (CPS) data obtained from the IPUMS-USA database (Sarah Flood and Warren (2017)) and task data from the US Occupational Information Network (O*NET). This is done in two steps: from the CPS I measure hours worked by employees in different occupations by year and industry; from O*NET I obtain measures of task inputs in each occupation. Multiplying year-industry-occupation hours by the corresponding measure of task inputs, I obtain a measure of worker task inputs in production. I now describe this procedure in detail.

The O*NET database consists of detailed data on a large number (approximately 1100) of individual occupations in the US economy. Each occupation is described by several numerical evaluations of the tasks performed by workers, as well as the the knowledge, skills, and abilities required to perform said tasks. To construct my baseline measure of task input in each occupation, I add the number of tasks reported as relevant for each occupation, weighted by each tasks relevance score. The score corresponds the fraction of workers in that occupation that reported that a given task is relevant to their job. At the level of detail at which I construct this measure, there are approximately 330 different tasks.

Using CPS data from 1972-2010, I construct measures of hours worked by year, industry and occupation. I multiply these occupation level variables by the corresponding task input measure from O*NET and sum over occupations to construct year-industry-task hour variables. The sum over tasks of this measure, divided by the total number of hours worked within each industry and the labor input quality from KLEMS, is my data counterpart to

$\Lambda_{j,t}$ as defined in equation 30.

$$\widehat{\Lambda}_{j,t} \equiv \sum_{\text{Tasks}} \sum_{\text{Occ}} \left(\frac{\text{Occupation Hours}_{j,t} \times \text{Occupation Task Input}}{\widehat{H}_{j,t} \widehat{N}_{j,t}} \right) \tag{35}$$

The final step is to construct a data counterpart to $\Delta \log M_{j.t}$, the log change in machine tasks. To do so, I designate a subset of the total tasks in the O*NET sample as automatable. I follow the extensive literature on automation and job polarization starting with Autor et al. (2003), in designating routine tasks as automatable. I denote the sum of this subset of tasks $\widehat{\Lambda}_{M,j,t}$, defined analogously to $\widehat{\Lambda}_{j,t}$:

$$\widehat{\Lambda}_{M,j,t} \equiv \sum_{\text{Auto Tasks}} \sum_{\text{Occ}} \left(\frac{\text{Occupation Hours}_{j,t} \times \text{Occupation Task Input}}{\widehat{H}_{j,t} \widehat{N}_{j,t}} \right) \tag{36}$$

To obtain the log change in tasks performed by machines $(\Delta \log \widehat{M}_{j,t})$ from the log change in automatable tasks $\Delta \log \widehat{\Lambda}_{M,j,t}$, I derive the following expression in Appendix B.1:

$$\Delta \log \widehat{M}_{j,t} = -\frac{\widehat{\Lambda}_{M,j,t-1}}{\widehat{\Lambda}_{j,t-1}} \frac{\widehat{\Lambda}_{j,t-1}}{\widehat{M}_{j,t-1}} \Delta \log \widehat{\Lambda}_{M,j,t} \tag{37}$$

With these expressions in hand, I solve for

$$\Delta \log \hat{a}_{h,j,t} = \frac{1}{\chi_{h,j,t}} \left(\Delta \log \hat{\Lambda}_{j,t} - \Delta \log \hat{M}_{j,t} - \chi_{w,j,t} \Delta \log \hat{a}_{w,j,t} \right)$$

which allows me to derive the model-implied labor share in each sector.

5 Automation and the Labor Share

In this section I describe the procedure used to fit the model to industry level data and present the results.

5.1 Fitting the model to data

In this section I explore the fit of model implied labor shares to industry labor share data from KLEMS. Given the data set $\left\{\hat{w}_{j,t},\ \widehat{H}_{j,t},\ \Delta\log\hat{\Lambda}_{j,t}\ ,\Delta\log\hat{M}_{j,t}\right\}_{t=1}^{T}$ for each sector j, two parameters fully determine the model-implied path of industry labor shares: the elasticity of substitution, σ_j, and the automation level of the lowest active firm at the start of the sample, $\hat{a}_{w,j,0}$. With these two parameters, the entire sequence of $\hat{a}_{h,j}$, $\hat{a}_{w,j}$ and hence $\widehat{LS}_j = \frac{1-\hat{a}_{h,j}}{1-\hat{a}_{w,j}}$ can be solved for each industry j as follows:

1. Solve for $\hat{a}_{h,j,0}$ using the labor share from the data at time zero and $\hat{a}_{w,j,0}$, $\hat{a}_{h,j,0} = 1 - LS_{j,0}\left(1-a_{w,j,0}\right)$
2. Solve for $\left\{\Delta\log\hat{a}_{w,j,t} = (1-\sigma)\frac{1-\hat{a}_{w,j,t-1}}{\hat{a}_{w,j,t-1}}\left(\Delta\log\hat{w}_{j,t} - \Delta\log\widehat{A}_{N,j,t}\right)\right\}_{t=1}^{T}$ starting from $\hat{a}_{w,j,0}$ and updating $\hat{a}_{w,j,t+1} = \hat{a}_{w,j,t}\times\exp\left(\Delta\log\hat{a}_{w,j,t}\right)$.
3. Iteratively solve for $\Delta\log\hat{a}_{h,j,t} = \frac{1}{\chi_{h,j,t}}\left(\Delta\log\hat{\Lambda}_{jt} - \Delta\log\hat{M}_{jt} - \chi_{w,j,t}\Delta\log\hat{a}_{w,j,t}\right)$ and update $\hat{a}_{h,j,t+1} = \hat{a}_{h,j,t}\times\exp\left(\Delta\log\hat{a}_{h,j,t}\right)$. The iteration is necessary because the coefficients $\chi_{h,j,t}$ and $\chi_{w,j,t}$ depend on $\hat{a}_{h,j,t-1}$ and $\hat{a}_{w,j,t-1}$.
4. Construct $\widehat{LS}_j = \frac{1-\hat{a}_{h,j}}{1-\hat{a}_{w,j}}$

To find the best fit of the model to the data, I numerically solve the following quadratic distance minimization problem between the model implied and KLEMS data industry labor share:

$$\min_{\hat{\sigma}_j,\hat{a}_{w,j,0}}\sum_{t=0}^{T}\left(\widehat{LS}_{j,t} - LS_{j,t}\right)^2 \tag{38}$$

The $\hat{\sigma}_j$ and $\hat{a}_{w,j,0}$ that minimize 38 are found by iterating over steps 1-4 above.

5.2 Results

In mapping the data to the model I am thinking of the economy as being on its balanced growth path at each point of observation. To focus on these low frequency movements, I filter the data using an HP filter with parameter 6.25 (the recommended parameter for annual data in Ravn and Uhlig (2002)).

Figure 2 plots the data and model-implied labor shares for 12 industries (comprising the business sector excluding finance, insurance and real estate) by value added in 1972. Table

3 shows the data and model implied change in the industry labor share, R^2, best fit $\hat{\sigma}$ and model-implied increase in $\hat{a}_h$ and $\hat{a}_w$ for those industries for which the model has a positive R^2. The fit of the model to the low frequency dynamics of industry-level labor shares varies significantly between industries. Of the 6 largest sectors by value added share in 1972, the model does reasonable well in fitting the low frequency trends in manufacturing, health and professional and business services, but poorly for retail and wholesale trade, construction and education. The model does not resort to extreme values of σ to fit the low frequency dynamics; for example the value of $\hat{\sigma}$ for manufacturing is 0.86; Oberfield and Raval (2014) estimate a σ in manufacturing around 0.75 in 2007. To form an idea of the overall fit of the model, Figure 3 plots the model implied aggregate labor share taking the value added of each sector from the KLEMS data.

Industry	ΔLS (data)	ΔLS (model)	R^2	$\hat{\sigma}$	$\Delta\hat{a}_h$	$\Delta\hat{a}_w$
Manufacturing	-0.18	-0.11	0.58	0.86	0.12	0.03
Health	0.08	0.09	0.81	0.93	-0.07	0.03
Prof. and Bus. Services	0.08	0.11	0.63	0.95	-0.05	0.02
Agriculture	0.05	0.08	0.71	0.84	-0.06	0.03
Information	-0.19	-0.15	0.63	0.91	0.18	0.05
Accommodation and Food	-0.12	-0.09	0.27	0.73	0.10	0

L-R: Change in LS in KLEMS data; change in LS in model; R^2 of fitted model; best fit σ; model implied change in a_h; model implied change in a_w. All data are fed into model after HP filtering with filter parameter 6.25.

Table 3: Results Fitting Model to KLEMS labor share data, 1972-2010

To illustrate the effect of automation on the industry and aggregate labor shares, Figure 4 plots the aggregate labor share and the labor share in manufacturing keeping $\hat{a}_{h,t}$ in manufacturing constant at its time zero value. The interpretation of this counterfactual is that the margin of automation that operates through the discovery of improved automation technologies is held constant, but the margin that operates through varying the lowest level of automation remains active. Within manufacturing, the labor share increases by approximately 0.02 in the counterfactual scenario, so is approximately 0.12 higher by 2010 than in the model baseline; the aggregate labor share is 0.02 higher in the counterfactual than in the baseline.

A concern with the analysis as it applies to manufacturing is that it is also by far the sector that is most exposed to import competition and offshoring of jobs. Gutierrez and

Philippon (2017) present evidence on the effect of competition from China on the US manufacturing industry. The authors show that mean import penetration of Chinese imports across manufacturing was close to zero in 1990, had risen to around 4% by 2000, and increased to 20% by 2015. Other measures of the effect of Chinese competition, such as entry and exit rates by exposure to Chinese competition, show a similar temporal pattern: a slow increase from a low base between 1990-2000, and a sharp acceleration thereafter. My results on the effect of automation on the manufacturing labor share (see top left panel of Figure 1) fail to match the sharp decline in the labor share that is evident in the KLEMS data post-2000, which may be due to the increased influence of offshoring and other effects of foreign competition.

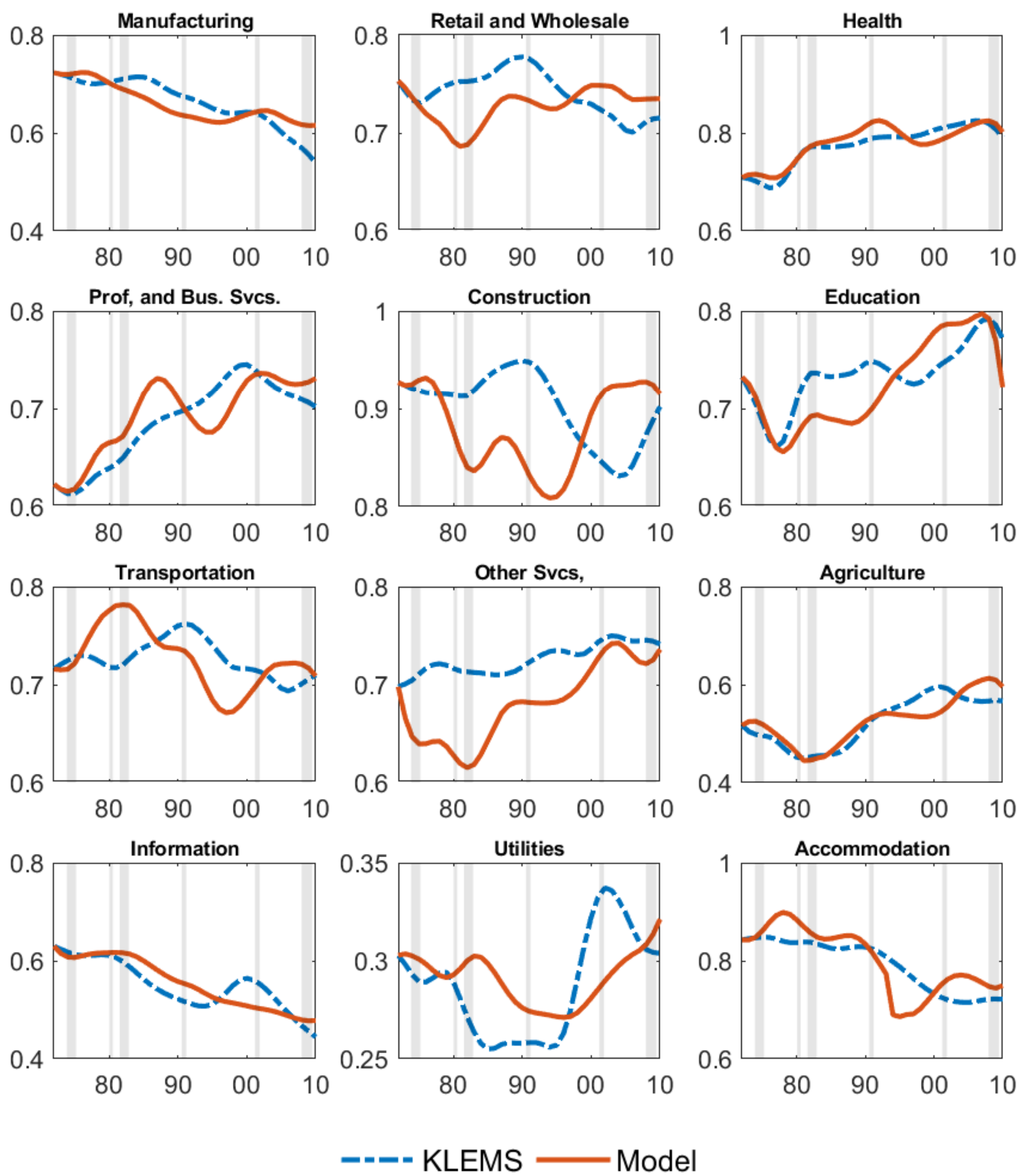
Overall, the empirical exercise provides suggestive evidence of an important role for automation in driving industry-level labor shares. One salient discrepancy between model and data is that the model fails to match most of the steep post-2000 drop in the aggregate labor share.

6 Conclusion

This paper develops a model of the aggregate production function in which the parameters of a CES production function are determined by moments of an underlying distribution of automation technologies. Automation in the model is a form of embodied technology which allows machines to replace workers in the production of tasks.

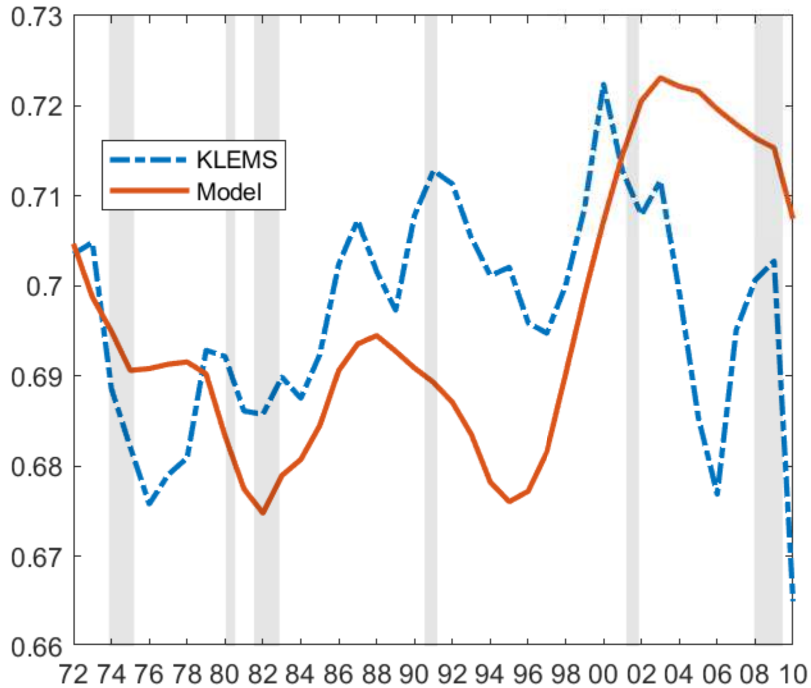
The balanced growth path of the model closely resembles that of a neo-classical growth model with a CES aggregate production function with sub-unitary elasticity of substitution and growth coming from labor augmenting technical progress, with one crucial difference: the labor share in the steady state is determined by the distribution of automation technology, and a balanced growth path exists for any strictly positive labor share. The model can rationalize trends in the labor share as arising from transitions between balanced growth paths, caused by periods of unbalanced growth between automation technology and human task inputs into production.

In the empirical exercise, the model is used to measure the effect of automation on industry-level labor shares in the US economy. To do so, I introduce a novel measure of aggregate task inputs into production, constructed using CPS data and O*NET. The



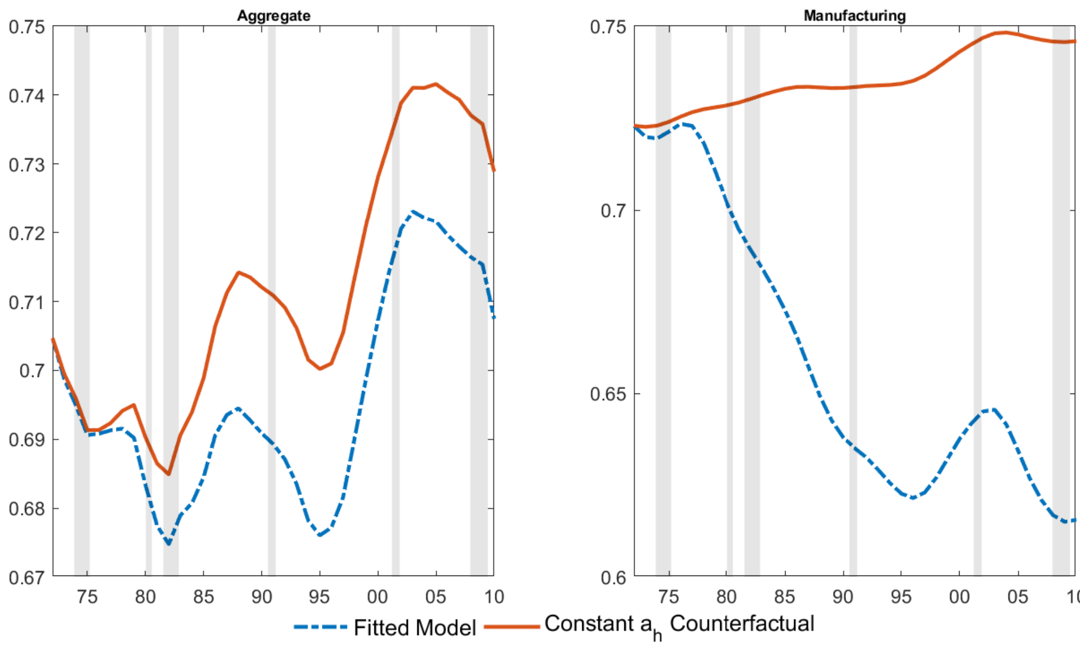
KLEMS data and model implied aggregate labor shares for 12 US industries

Figure 2: KLEMS and model implied labor shares for 12 US Industries, 1972-2010



KLEMS data and model implied aggregate labor shares. The model aggregate is constructed using value added shares from the data to construct the aggregate from the model-implied industry level shares.

Figure 3: KLEMS and model implied aggregate labor share, 1972-2008



Model implied labor shares of the fitted model and a counterfactual in which the a_h in manufacturing is held constant, but a_w follows its fitted path. The interpretation of the counterfactual is that the technological innovation margin of automation in manufacturing is held constant but the endogenous scrapping margin follows its fitted path. The left panel plots the aggregate labor share and the right panel the labor share in manufacturing.

Figure 4: Model implied labor share and model counterfactual with constant a_h in manufacturing

model fits the HP-filtered trend in industry labor shares reasonably well for some important industries, and the fitted model implies that automation was a significant force driving fluctuations in the labor share in data from 1972-2010.

References

Acemoglu, D. and D. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Hanbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, North Holland, vol. 4b.

Acemoglu, D. and P. Restrepo (2016): "The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment," Working Paper 22252, National Bureau of Economic Research.

Aghion, P., B. F. Jones, and C. I. Jones (2017): "Artificial Intelligence and Economic Growth," Working Paper 23928, National Bureau of Economic Research.

Akinsete, A., F. Famoye, and C. Lee (2008): "The beta-Pareto distribution," *Statistics*, 42, 547–563.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. V. Reenen (2017): "The Fall of the Labor Share and the Rise of Superstar Firms," Working Paper 23396, National Bureau of Economic Research.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118, 1279–1333.

Barkai, S. (2017): "Declining Labor and Capital Shares," Working Paper.

Baumol, W. J. (1967): "Macroeconomics of Unbalanced Growth: The Anatomy of the Urban Crisis," *American Economic Review*, 57, 415–426.

Becker, G. S. and K. M. Murphy (1992): "The Division of Labor, Coordination Costs, and Knowledge*," *The Quarterly Journal of Economics*, 107, 1137–1160.

Benhabib, J., J. Perla, and C. Tonetti (2017): "Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier," Working Paper 23095, National Bureau of Economic Research.

Caballero, R. J., E. Farhi, and P.-O. Gourinchas (2017): "Rents, Technical Change, and Risk Premia: Accounting for Secular Trends in Interest Rates, Returns on Capital, Earning Yields, and Factor Shares," Working Paper 23127, National Bureau of Economic Research.

Chirinko, R. S. (2008): "Sigma: The long and short of it," *Journal of Macroeconomics*, 30, 671 – 686, the CES Production Function in the Theory and Empirics of Economic Growth.

de La Grandville, O. (2016): *Economic Growth: A Unified Approach*, Cambridge University Press, 2 ed.

De Loecker, J. and J. Eeckhout (2017): “The Rise of Market Power and the Macroeconomic Implications,” Working Paper 23687, National Bureau of Economic Research.

Elsby, M., B. Hobijn, and A. Sahin (2013): “The decline of the U.S. labor share,” Working Paper Series 2013-27, Federal Reserve Bank of San Francisco.

Garicano, L. (2000): “Hierarchies and the Organization of Knowledge in Production,” *Journal of Political Economy*, 108, 874–904.

Garicano, L. and E. Rossi-Hansberg (2006): “Organization and Inequality in a Knowledge Economy*,” *The Quarterly Journal of Economics*, 121, 1383–1435.

Geerolf, F. (2017): “A Theory of Pareto Distributions,” Working Paper.

Gilchrist, S. and J. C. Williams (2000): “Putty-Clay and Investment: A Business Cycle Analysis,” *Journal of Political Economy*, 108, 928–960.

Gutierrez, G. (2017): “Declining Competition and Labor Shares in the US,” *Working Paper.*

Gutierrez, G. and T. Philippon (2017): “Declining Competition and Investment in the US,” Working Paper.

Hartman-Glaser, B., H. Lustig, and M. Z. Xiaolan (2017): “Capital Share Dynamics When Firms Insure Workers,” Research Paper 16-35, Stanford University Graduate School of Business.

Hemous, D. and M. Olsen (2016): “The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality,” Working Paper.

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

Johansen, L. (1959): “Substitution versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis,” *Econometrica: Journal of the Econometric Society*, 27, 157–176.

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

Jones, C. I. and D. Scrimgeour (2008): “A New Proof of Uzawa’s Steady-State Growth Theorem,” *The Review of Economics and Statistics*, 90, 180–182.

Jorgenson, D. W., M. S. Ho, and J. Samuels (2012): "A Prototype Industry-Level Production Account for the United States, 1947-2010," Second World KLEMS Conference, Harvard University.

Jovanovic, B. and Y. Yatsenko (2012): "Investment in vintage capital," *Journal of Economic Theory*, 147, 551 – 569, issue in honor of David Cass.

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

Kehrig, M. and N. Vincent (2017): "Growing Productivity Without Growing Wages: The Micro-Level Anatomy of the Aggregate Labor Share Decline," *Working Paper*.

King, R. G., C. I. Plosser, and S. T. Rebelo (1988): "Production, Growth and Business Cycles: The Basic Neoclassical Model," *Journal of Monetary Economics*, 21, 195–232.

Koh, D., R. Santaeulalia-Llopis, and Y. Zheng (2016): "Labor Share Decline and Intellectual Property Products Capital," Working Paper.

Lagos, R. (2006): "A Model of TFP," *The Review of Economic Studies*, 73, 983.

Lawrence, R. Z. (2015): "Recent Declines in Labor's Share in US Income: A Preliminary Neoclassical Account," Working Paper 21296, National Bureau of Economic Research.

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

Lorenzutti, R., D. Duarte, and M. Azevedo (2012): "The Beta Truncated Pareto Distribution," Mimeo.

Mangin, S. (2015): "How to Build a Production Function," Working Paper.

Oberfield, E. and D. Raval (2014): "Micro Data and Macro Technology," Working Paper 20452, National Bureau of Economic Research.

Ravn, M. O. and H. Uhlig (2002): "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *The Review of Economics and Statistics*, 84, 371–376.

Rognlie, M. (2015): "Deciphering the fall and rise in the net capital share," *Brookings papers on economic activity*, 6.

Sarah Flood, Miriam King, S. R. and J. R. Warren (2017): *Integrated Public Use Microdata Series, Current Population Survey: Version 5.0 (dataset)*, Minneapolis: University of Minnesota.

Sato, K. (1969): "Micro and macro constant-elasticity-of-substitution production functions in a multifirm industry," *Journal of Economic Theory*, 1, 438 – 453.

——— (1975): *Production functions and aggregation*, vol. 90 of *Contributions to Economic Analysis*, North-Holland/American Elsevier.

Susskind, D. (2017): "A Model of Technological Unemployment," Working Paper 819, Oxford University.

Uzawa, H. (1961): "Neutral Inventions and the Stability of Growth Equilibrium," *The Review of Economic Studies*, 28, 117–124.

Zeira, J. (1998): "Workers, Machines, and Economic Growth," *The Quarterly Journal of Economics*, 113, 1091–1117.

A Proofs

A.1 Proof of Lemma 1

Start with the CDF $G(z)$.

$$G(z) = 1 - \left(\frac{\left(\frac{z_\ell}{z}\right)^{\frac{1}{\gamma}} - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}}{1 - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}} \right)^{\rho}$$

The random variable $\tilde{u} = G(z)$ is uniformly distributed in $[0,1]$ (by the probability integral transformation).

$$\begin{aligned}
\tilde{u} &= 1 - \left(\frac{\left(\frac{z_\ell}{z}\right)^{\frac{1}{\gamma}} - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}}{1 - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}} \right)^{\rho} \\
u = 1 - \tilde{u} &= \left(\frac{\left(\frac{z_\ell}{z}\right)^{\frac{1}{\gamma}} - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}}{1 - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}} \right)^{\rho}
\end{aligned}$$

If $\tilde{u} \sim U[0,1]$, then $u = 1 - \tilde{u}$ is also $u \sim U[0,1]$. Now, express z as a function of u:

$$\begin{aligned}
u^{\frac{1}{\rho}} &= \frac{\left(\frac{z_\ell}{z}\right)^{\frac{1}{\gamma}} - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}}{1 - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}} \\
\left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}} + \left(1 - \left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}\right) u^{\frac{1}{\rho}} &= \left(\frac{z_\ell}{z}\right)^{\frac{1}{\gamma}} \\
\left(\left(\frac{1}{z_h}\right)^{\frac{1}{\gamma}} + \left(\left(\frac{1}{z_\ell}\right)^{\frac{1}{\gamma}} - \left(\frac{1}{z_h}\right)^{\frac{1}{\gamma}} \right) u^{\frac{1}{\rho}} \right)^{-\gamma} &= z
\end{aligned}$$

By the probability integral transform, the random variable $\tilde{\chi} = u^{\frac{1}{\rho}}$ is distributed $B(\rho, 1)$, $\tilde{\chi} \in [0,1]$. It follows that the random variable $\chi = z_h^{-\frac{1}{\gamma}} + \left(z_\ell^{-\frac{1}{\gamma}} - z_h^{-\frac{1}{\gamma}} \right) \tilde{\chi}$ is distributed $B(\rho, 1)$ with support $\chi \in \left[z_h^{-\frac{1}{\gamma}}, z_\ell^{-\frac{1}{\gamma}} \right]$. So $z = \chi^{-\gamma}$ has distribution$G(z)$. Now, let $a \equiv 1 - \chi$.

Since $\chi \sim B\left(\rho, 1; z_h^{-\frac{1}{\gamma}}, z_\ell^{-\frac{1}{\gamma}}\right)$, $a \sim B(1, \rho; a_\ell, a_h)$, with $a_\ell \equiv 1 - z_\ell^{-\frac{1}{\gamma}}$ and $a_h \equiv 1 - z_h^{-\frac{1}{\gamma}}$. Finally, let $\tilde{a} = qa$. Then, $\tilde{a} \sim B(1, \rho; \tilde{a}_\ell, \tilde{a}_h)$, with $a_\ell = \frac{\tilde{a}_\ell}{q}$ and $a_h = \frac{\tilde{a}_h}{q}$[9], and

$$z = \left(\frac{q}{q - \tilde{a}}\right)^\gamma = \left(\frac{1}{1-a}\right)^\gamma$$

A.2 Proof of Proposition 1

I first show how to derive the distribution $G(z)$ from the desired CES production function.

A.2.1 From CES Production Function to Distribution

Want a production function of the form:

$$Y = \left(\kappa K^{\frac{\sigma-1}{\sigma}} + \nu N^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

Dividing both sides by K,

$$\frac{Y}{K} = \left(\kappa + \nu \left(\frac{N}{K}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

Define $y \equiv \frac{Y}{K}$ and $n \equiv \frac{N}{K}$.

$$y = \left(\kappa + \nu n^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{39}$$

Express y and n as functions of $\tilde{w}$:

[9] Proof:

$$f(\chi) = \frac{\rho\left(\chi - H^{-\frac{1}{\gamma}}\right)^{\rho-1}}{\left(L^{-\frac{1}{\gamma}} - H^{-\frac{1}{\gamma}}\right)^\rho}$$

Substitute in $a = 1 - \chi$

$$f(x) = \frac{\rho\left(1 - a - H^{-\frac{1}{\gamma}}\right)^{\rho-1}}{\left(L^{-\frac{1}{\gamma}} - H^{-\frac{1}{\gamma}}\right)^\rho} = \frac{\rho\left((1-a_h) - a\right)^{\rho-1}}{\left((1-a_h) - \left(1 - L^{-\frac{1}{\gamma}}\right)\right)^\rho} = \frac{\rho\left(a_h - a\right)^{\rho-1}}{\left(a_h - a_\ell\right)^\rho}$$

So $x \sim B(1, \rho; \ell, h, 1 - a_h)$. Now, let $\tilde{a} = qx$.

$$f(\tilde{a}) = \frac{1}{q}\frac{\rho\left(a_h - \frac{\tilde{a}}{q}\right)^{\rho-1}}{\left(a_h - a_\ell\right)^\rho} = \frac{1}{q^\rho}\frac{\rho\left(qa_h - \tilde{a}\right)^{\rho-1}}{\left(a_h - a_\ell\right)^\rho} = \frac{\rho\left(\tilde{a}_h - \tilde{a}\right)^{\rho-1}}{\left(\tilde{a}_h - \tilde{a}_\ell\right)^\rho}$$

$$y\left(\tilde{w}\right)^{\frac{\sigma-1}{\sigma}}=\kappa+\nu n\left(\tilde{w}\right)^{\frac{\sigma-1}{\sigma}} \tag{40}$$

Total derivative w.r.t. $\tilde{w}$,

$$\left(\frac{\sigma-1}{\sigma}\right)y\left(\tilde{w}\right)^{-\frac{1}{\sigma}}\frac{\partial y}{\partial\tilde{w}}=\left(\frac{\sigma-1}{\sigma}\right)\nu n\left(\tilde{w}\right)^{-\frac{1}{\sigma}}\frac{\partial n}{\partial\tilde{w}}$$

It is convenient to define $y\equiv\frac{Y}{K}$ and $n\equiv\frac{N}{K}$:

$$y=\int_{\tilde{w}}^{H}zg\left(z\right)\,\mathrm{d}z \tag{41}$$

$$n=\frac{1}{q}\left(1-G\left(\tilde{w}\right)\right) \tag{42}$$

From equations (41) and (42),

$$\frac{\partial y}{\partial\tilde{w}}=-\tilde{w}g\left(\tilde{w}\right)$$

$$\frac{\partial n}{\partial\tilde{w}}=-\frac{1}{q}g\left(\tilde{w}\right)$$

Substituting,

$$\begin{aligned}
-y\left(\tilde{w}\right)^{-\frac{1}{\sigma}}\tilde{w}g\left(\tilde{w}\right) &= -\nu n\left(\tilde{w}\right)^{-\frac{1}{\sigma}}\frac{1}{q}g\left(\tilde{w}\right)\\
y\left(\tilde{w}\right)^{-\frac{1}{\sigma}} &= \left(\frac{\nu}{q\tilde{w}}\right)n\left(\tilde{w}\right)^{-\frac{1}{\sigma}}\\
y\left(\tilde{w}\right)^{-1} &= \left(\frac{\nu}{H\cdot\tilde{w}}\right)^{\sigma}n\left(\tilde{w}\right)^{-1}\\
y\left(\tilde{w}\right) &= \left(\frac{H\cdot\tilde{w}}{\nu}\right)^{\sigma}n\left(\tilde{w}\right)
\end{aligned}$$

Substituting into equation (40):

$$\begin{aligned}
\left(\left(\frac{H\cdot\tilde{w}}{\nu}\right)^{\sigma}n\left(\tilde{w}\right)\right)^{\frac{\sigma-1}{\sigma}} &= \kappa+\nu n\left(\tilde{w}\right)^{\frac{\sigma-1}{\sigma}}\\
\left(\frac{H\cdot\tilde{w}}{\nu}\right)^{\sigma-1}n\left(\tilde{w}\right)^{\frac{\sigma-1}{\sigma}} &= \kappa+\nu n\left(\tilde{w}\right)^{\frac{\sigma-1}{\sigma}}\\
\left(\left(\frac{H\cdot\tilde{w}}{\nu}\right)^{\sigma-1}-\nu\right)n\left(\tilde{w}\right)^{\frac{\sigma-1}{\sigma}} &= \kappa\\
n\left(\tilde{w}\right)^{\frac{\sigma-1}{\sigma}} &= \kappa\left(\left(\frac{H\cdot\tilde{w}}{\nu}\right)^{\sigma-1}-\nu\right)^{-1}\\
n\left(\tilde{w}\right) &= \kappa^{\frac{\sigma}{\sigma-1}}\left(\left(\frac{H\cdot\tilde{w}}{\nu}\right)^{\sigma-1}-\nu\right)^{-\frac{\sigma}{\sigma-1}}\\
&= \kappa^{\frac{\sigma}{\sigma-1}}\left(\left(\frac{H\cdot\tilde{w}}{\nu}\right)^{\sigma-1}-\nu\right)^{\frac{\sigma}{1-\sigma}}
\end{aligned}$$

Substitute in from equation (42),

$$\begin{aligned}
\frac{1}{H}\left(1-G\left(z\right)\right) &= \kappa^{\frac{\sigma}{\sigma-1}}\left(\left(\frac{H\cdot z}{\nu}\right)^{\sigma-1}-\nu\right)^{\frac{\sigma}{1-\sigma}}\\
G\left(z\right) &= 1-H\cdot\kappa^{\frac{\sigma}{\sigma-1}}\left(\left(\frac{H\cdot z}{\nu}\right)^{\sigma-1}-\nu\right)^{\frac{\sigma}{1-\sigma}}
\end{aligned}$$

By definition, G is a CDF and $z\in[z_\ell,z_h]$, so $G\left(H\right)=1$ and $G\left(L\right)=0$. To solve for κ and ν, evaluate at the boundaries:

$$\begin{aligned}
1 &= G\left(z_h\right)\\
0 &= H\cdot\kappa\left(\left(\frac{H\cdot z_h}{\nu}\right)^{\sigma-1}-\nu\right)^{\frac{\sigma}{1-\sigma}}\\
0 &= \left(\frac{H\cdot z_h}{\nu}\right)^{\sigma-1}-\nu\\
\nu^{\sigma} &= \left(H\cdot z_h\right)^{\sigma-1}
\end{aligned}$$

$$\nu = \left(H \cdot z_h\right)^{\frac{\sigma-1}{\sigma}} \tag{43}$$

Substitute in,

$$\begin{aligned}
G\left(z\right) &= 1 - H \cdot \kappa^{\frac{\sigma-1}{\sigma}} \left(\left(\frac{H \cdot z}{\left(H \cdot z_h\right)^{\frac{\sigma-1}{\sigma}}} \right)^{\sigma-1} - \left(H \cdot z_h\right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{1-\sigma}} \\
&= 1 - H \cdot \kappa^{\frac{\sigma-1}{\sigma}} \left(\left(H \cdot z_h\right)^{\frac{\sigma-1}{\sigma}} \left(\left(H \cdot z_h\right)^{1-\sigma} \left(H \cdot z\right)^{\sigma-1} - 1 \right) \right)^{\frac{\sigma}{1-\sigma}} \\
&= 1 - H \cdot \kappa^{\frac{\sigma-1}{\sigma}} \left(H \cdot z_h\right)^{-1} \left(\left(\left(H \cdot z_h\right)^{1-\sigma} \left(H \cdot z\right)^{\sigma-1} - 1 \right) \right)^{\frac{\sigma}{1-\sigma}} \\
&= 1 - \kappa^{\frac{\sigma-1}{\sigma}} z_h^{-1} \left(\left(\frac{z_h}{z} \right)^{1-\sigma} - 1 \right)^{\frac{\sigma}{1-\sigma}}
\end{aligned}$$

Evaluate $G\left(L\right)$:

$$\begin{aligned}
0 &= G\left(z_\ell\right) \\
1 &= \kappa^{\frac{\sigma}{\sigma-1}} z_h^{-1} \left(\left(\frac{z_h}{z_\ell} \right)^{1-\sigma} - 1 \right)^{\frac{\sigma}{1-\sigma}} \\
\kappa^{\frac{\sigma}{\sigma-1}} &= z_h \left(\left(\frac{z_h}{z_\ell} \right)^{1-\sigma} - 1 \right)^{\frac{\sigma}{\sigma-1}}
\end{aligned}$$

$$\kappa = z_h^{\frac{\sigma-1}{\sigma}} \left(\left(\frac{z_h}{z_\ell} \right)^{1-\sigma} - 1 \right) \tag{44}$$

Substitute in,

$$
\begin{aligned}
G(z) &= 1-\left(z_h^{\frac{\sigma-1}{\sigma}}\left(\left(\frac{z_h}{z_\ell}\right)^{1-\sigma}-1\right)\right)^{\frac{\sigma}{\sigma-1}} z_h^{-1}\left(\left(\frac{z_h}{z}\right)^{1-\sigma}-1\right)^{\frac{\sigma}{1-\sigma}} \\
&= 1-z_h\left(\left(\frac{z_h}{z_\ell}\right)^{1-\sigma}-1\right)^{\frac{\sigma}{\sigma-1}} z_h^{-1}\left(\left(\frac{z_h}{z}\right)^{1-\sigma}-1\right)^{\frac{\sigma}{1-\sigma}} \\
&= 1-\frac{1}{\left(\left(\frac{z_h}{z_\ell}\right)^{1-\sigma}-1\right)^{\frac{\sigma}{1-\sigma}}}\left(\left(\frac{z_h}{z}\right)^{1-\sigma}-1\right)^{\frac{\sigma}{1-\sigma}} \\
&= 1-\left(\frac{\left(\frac{z_h}{z}\right)^{1-\sigma}-1}{\left(\frac{z_h}{z_\ell}\right)^{1-\sigma}-1}\right)^{\frac{\sigma}{1-\sigma}} \\
&= 1-\left(\frac{\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}\left(\left(\frac{z_h}{z}\right)^{1-\sigma}-1\right)}{\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}\left(\left(\frac{z_h}{z_\ell}\right)^{1-\sigma}-1\right)}\right)^{\frac{\sigma}{1-\sigma}} \\
&= 1-\left(\frac{\left(\frac{z_\ell}{z}\right)^{1-\sigma}-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{1-\sigma}}
\end{aligned}
$$

Rearrange:

$$
G(z)=1-\left(1-\frac{1-\left(\frac{z_\ell}{z}\right)^{1-\sigma}}{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{1-\sigma}}
$$

With $\rho=\frac{\sigma}{1-\sigma}$ and $\gamma=\frac{\sigma}{1-\sigma}+1=\frac{1}{1-\sigma}$,

$$
G(z)=1-\left(1-\frac{1-\left(\frac{z_\ell}{z}\right)^{\frac{1}{\gamma}}}{1-\left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}}\right)^{\rho}
$$

A.2.2 The CES production function in normalized form

Substitute ν and κ from equations (43) and (44) into the CES production function (40):

$$y = \left(\kappa + \nu n^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$y = \left(z_h^{\frac{\sigma-1}{\sigma}} \left(\left(\frac{z_h}{z_\ell}\right)^{1-\sigma} - 1\right) + (z_h H)^{\frac{\sigma-1}{\sigma}} n^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$y = z_h \left(\left(\left(\frac{z_h}{z_\ell}\right)^{1-\sigma} - 1\right) + (H \cdot n)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

Define:

$$\alpha \equiv \frac{\left(\frac{z_h}{z_\ell}\right)^{1-\sigma} - 1}{\left(\frac{z_h}{z_\ell}\right)^{1-\sigma} - 1 + 1} = \frac{\left(\frac{z_h}{z_\ell}\right)^{1-\sigma} - 1}{\left(\frac{z_h}{z_\ell}\right)^{1-\sigma}} = 1 - \left(\frac{z_\ell}{z_h}\right)^{1-\sigma}$$

$$1 - \alpha \equiv 1 - \left(1 - \left(\frac{z_\ell}{z_h}\right)^{1-\sigma}\right) = \left(\frac{z_\ell}{z_h}\right)^{1-\sigma}$$

Substitute:

$$y = z_h \left(\left(\frac{z_h}{z_\ell}\right)^{1-\sigma}\right)^{\frac{\sigma}{\sigma-1}} \left(\alpha + (1-\alpha)(H \cdot n)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$y = z_h \left(\frac{z_h}{z_\ell}\right)^{-\sigma} \left(\alpha + (1-\alpha)(H \cdot n)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$y = \left((z_h)^{1-\sigma} (z_\ell)^{\sigma}\right) \left(\alpha + (1-\alpha)(H \cdot n)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

Multiply both sides by K,

$$Y = \left((z_h)^{1-\sigma} (z_\ell)^{\sigma}\right) \left(\alpha K^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(H \cdot n)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$Y = A \left(\alpha K^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(H \cdot n)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

Where $A = (z_h)^{1-\sigma} (z_\ell)^{\sigma}$; $\alpha = 1 - \left(\frac{z_\ell}{z_h}\right)^{1-\sigma}$. Now, substitute in the definitions of H and L.

A.2.3 A, α in terms of distribution of a

$$
\begin{aligned}
A &= (z_h)^{1-\sigma}(z_\ell)^{\sigma} \\
&= \left(\left(\frac{1}{1-a_h}\right)^{\frac{1}{1-\sigma}}\right)^{1-\sigma}\left(\left(\frac{1}{1-a_\ell}\right)^{\frac{1}{1-\sigma}}\right)^{\sigma} \\
&= \frac{1}{1-a_h}\left(\frac{1}{1-a_\ell}\right)^{\frac{\sigma}{1-\sigma}} \\
&= \frac{1}{1-a_h}(1-a_\ell)^{\frac{\sigma}{\sigma-1}} \\
&= (1-a_\ell)^{\frac{1}{\sigma-1}}\left(\frac{1-a_\ell}{1-a_h}\right)
\end{aligned}
$$

$$
\begin{aligned}
\alpha &= 1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma} \\
&= 1-\left(\frac{\left(\frac{1}{1-a_\ell}\right)^{\frac{1}{1-\sigma}}}{\left(\frac{1}{1-a_h}\right)}\right)^{1-\sigma} \\
&= 1-\frac{\left(\frac{1}{1-a_\ell}\right)}{\left(\frac{1}{1-a_h}\right)} \\
&= 1-\frac{1-a_h}{1-a_\ell} \\
&= \frac{1-a_\ell-1+a_h}{1-a_\ell} \\
&= \frac{a_h-a_\ell}{1-a_\ell}
\end{aligned}
$$

A.2.4 Labor share

Labor share of the CES production function:

$$
LS = (1-\alpha)\left(\frac{Y}{A\cdot H\cdot N}\right)^{\frac{1-\sigma}{\sigma}}
$$

Substituting in for A and α gives:

$$LS = (1 - a_h)\left(\frac{w}{H}\right)^{1-\sigma}$$

A.2.5 From Distribution to CES Production Function

$$\begin{aligned}
y &= \int_{w/H}^{z_h} z g(z)\ \mathrm{d}z \\
&= \int_{w/H}^{z_h} \sigma \frac{\left(\frac{z_\ell}{z}\right)^{1-\sigma}}{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}} \left(1 - \frac{1-\left(\frac{z_\ell}{z}\right)^{1-\sigma}}{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{1-\sigma}-1} \mathrm{d}z \\
&= \left(\frac{\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{1-\sigma}} \left(0 - z_h\left(1-\left(\frac{w/H}{z_h}\right)^{1-\sigma}\right)^{\frac{\sigma}{1-\sigma}}\right) \\
&= \left(\frac{\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{1-\sigma}} \left(z_h\left(1-\left(\frac{w/H}{z_h}\right)^{1-\sigma}\right)^{\frac{\sigma}{1-\sigma}}\right) \\
&= (z_h)^{1-\sigma}(z_\ell)^{\sigma}\left(\frac{1-\left(\frac{w/H}{z_h}\right)^{1-\sigma}}{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{1-\sigma}} \\
&= (z_h)^{1-\sigma}(z_\ell)^{\sigma}\left(\frac{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}{1-\left(\frac{w/H}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{\sigma-1}}
\end{aligned}$$

Note

$$
\begin{aligned}
1-G\left(\tilde{w}\right) &= \left(\frac{\left(\frac{z_\ell}{\tilde{w}}\right)^{1-\sigma}-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{1-\sigma}} \\
\left(H\cdot n\right)^{\frac{1-\sigma}{\sigma}} &= \frac{\left(\frac{z_h}{w/H}\right)^{1-\sigma}\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}-\left(1-\alpha\right)}{\alpha} \\
\left(H\cdot n\right)^{\frac{1-\sigma}{\sigma}} &= \frac{\left(\frac{z_h}{w/H}\right)^{1-\sigma}\left(1-\alpha\right)-\left(1-\alpha\right)}{\alpha} \\
\frac{\alpha\left(H\cdot n\right)^{\frac{1-\sigma}{\sigma}}+\left(1-\alpha\right)}{1-\alpha} &= \left(\frac{z_h}{w/H}\right)^{1-\sigma} \\
\left(\frac{z_h}{w/H}\right)^{1-\sigma} &= 1+\frac{\alpha}{1-\alpha}\left(H\cdot n\right)^{\frac{1-\sigma}{\sigma}} \\
\left(\frac{z_h}{w/H}\right)^{1-\sigma} &= \frac{\alpha}{\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}}+1 \\
&= \frac{\alpha+\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}}{\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}} \\
\left(\frac{w/H}{z_h}\right)^{1-\sigma} &= \frac{\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}}{\alpha+\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}} \\
1-\left(\frac{w/H}{z_h}\right)^{1-\sigma} &= 1-\frac{\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}}{\alpha+\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}} \\
&= \frac{\alpha+\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}-\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}}{\alpha+\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}} \\
&= \frac{\alpha}{\alpha+\left(1-\alpha\right)\left(H\cdot n\right)^{\frac{\sigma-1}{\sigma}}}
\end{aligned}
$$

Substitute in

$$
\begin{aligned}
y &= (z_h)^{1-\sigma}(z_\ell)^{\sigma}\left(\frac{1-\left(\frac{z_\ell}{z_h}\right)^{1-\sigma}}{1-\left(\frac{w/H}{z_h}\right)^{1-\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \\
&= A\left(\frac{\alpha}{\frac{\alpha}{\alpha+(1-\alpha)(H\cdot n)^{\frac{\sigma-1}{\sigma}}}}\right)^{\frac{\sigma}{\sigma-1}} \\
&= A\left(\alpha+(1-\alpha)(H\cdot N)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}
\end{aligned}
$$

Multiply both sides by K:

$$
Y = A\left(\alpha K^{\frac{\sigma-1}{\sigma}}+(1-\alpha)(H\cdot N)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}
$$

B Derivations

B.1 Task production

B.1.1 Aggregate and machine task production

Worker task input λ_i at firm i:

$$
\begin{aligned}
\lambda_i = (q-\tilde{a}_i)\,y_i &= q(1-a_i)\,y_i \\
&= q(1-a_i)\,z(a_i)\,k(a_i) \\
&= qK(1-a_i)\,z(a_i)\,f(a_i) \\
&= q\frac{HN}{1-F(a_w)}(1-a_i)\,z(a_i)\,f(a_i) \\
&= qHN(1-a_i)\left(\frac{1}{1-a_i}\right)^{\rho+1}\frac{f(a_i)}{1-F(a_w)} \\
&= qHN\left(\frac{1}{1-a_i}\right)^{\rho}\frac{f(a_i)}{1-F(a_w)} \\
&= qHN\left(\frac{1}{1-a_i}\right)^{\rho}\frac{(a_h-a)^{\rho-1}}{(a_h-a_w)^{\rho}}
\end{aligned}
$$

Where I have used $k(a)\equiv Kf(a)$, $K=\frac{HN}{1-G\left(\frac{w}{H}\right)}=\frac{HN}{1-F(a_w)}$, $F(a)=\left(\frac{a_h-a}{a_h-a_\ell}\right)^{\rho}$, $z(a)=$

$\left(\frac{1}{1-a}\right)^{\rho+1}$

$$\Lambda = \frac{\int \lambda_i \text{ d}i}{HN} = q \int_{a_w}^{a_h} \left(\frac{1}{1-a}\right)^{\rho} \frac{(a_h - a)^{\rho-1}}{(a_h - a_w)^{\rho}} \text{ d}a$$

Noting that $\int_{a_w}^{a_h} \left(\frac{1}{1-a}\right)^{\rho} \frac{(a_h-a)^{\rho-1}}{(a_h-a_w)^{\rho}} \text{ d}a = \frac{1}{1-F(a_w)} \int_{a_w}^{a_h} (1-a) \text{ dF}a = \mathbb{E}_F\left(1-a|a > a_w\right)$, it follows that:

$$\Lambda = q\mathbb{E}_F\left[1-a|a > a_w\right]$$

Similarly, machine task input μ_i at firm i:

$$\begin{aligned} \mu_i &= \tilde{a}_i y_i = q a_i y_i \\ &= qHNa_i \left(\frac{1}{1-a_i}\right)^{\rho+1} \frac{(a_h - a)^{\rho-1}}{(a_h - a_w)^{\rho}} \end{aligned}$$

$$M = \frac{\int \mu_i \text{ d}i}{HN} = q \int_{a_w}^{a_h} a \left(\frac{1}{1-a}\right)^{\rho+1} \frac{(a_h - a)^{\rho-1}}{(a_h - a_w)^{\rho}} \text{ d}a$$

And as above,

$$M = q\mathbb{E}_F\left[a|a > a_w\right]$$

The ratio of aggregate worker to machine task inputs:

$$\frac{\Lambda}{M} = \frac{\int_{a_w}^{a_h} \left(\frac{1}{1-a}\right)^{\rho} (a_h - a)^{\rho-1} \text{ d}a}{\int_{a_w}^{a_h} a \left(\frac{1}{1-a}\right)^{\rho+1} (a_h - a)^{\rho-1} \text{ d}a} \tag{45}$$

Taking a first order expansion around q, a_h and a_w,

$$\Delta \log \Lambda = \Delta \log q + \chi_{\Lambda,h} \Delta \log a_h + \chi_{\Lambda,w} \Delta \log a_w \tag{46}$$

$$\Delta \log M = \Delta \log q + \chi_{M,h} \Delta \log h + \chi_{M,w} \Delta \log a_w \tag{47}$$

Subtracting 47 from 46:

$$\begin{aligned} \Delta \log \Lambda - \Delta \log M &= \left(\chi_{\Lambda,h} - \chi_{M,h}\right) \Delta \log a_h + \left(\chi_{\Lambda,w} - \chi_{M,w}\right) \Delta \log a_w \\ &= \chi_{h,t} \Delta \log a_{h,t} + \chi_{w,t} \Delta \log a_{w,t} \end{aligned}$$

B.1.2 Measuring change in machine tasks $\Delta \log M$ in the data

To approximate $\Delta \log M$ in the data, I designate a subset of tasks as automatable. I denote the sum of these tasks at time t as $\widehat{\Lambda}_{M,t}$, and I assume that the level change in machine tasks is equal to the (negative of) the level change change in automatable human tasks:

$$\widehat{M}_t - \widehat{M}_{t-1} = -\left(\widehat{\Lambda}_{M,t} - \widehat{\Lambda}_{M,t-1}\right)$$

Some algebra to get percentage changes in $\widehat{M}$ and $\widehat{\Lambda}$,

$$\frac{\widehat{M}_t - \widehat{M}_{t-1}}{\widehat{M}_{t-1}} = -\frac{\widehat{\Lambda}_{M,t-1}}{\widehat{M}_{t-1}}\frac{\widehat{\Lambda}_{M,t} - \widehat{\Lambda}_{M,t-1}}{\widehat{\Lambda}_{M,t-1}}$$

Approximating percentage changes with log differences,

$$\Delta\log\widehat{M}_t = -\frac{\widehat{\Lambda}_{M,t-1}}{\widehat{M}_{t-1}}\Delta\log\widehat{\Lambda}_{M,t}$$

And finally, substituting $\frac{\widehat{\Lambda}_{M,t-1}}{\widehat{M}_{t-1}} = \frac{\widehat{\Lambda}_{M,t-1}}{\widehat{\Lambda}_{t-1}}\frac{\widehat{\Lambda}_{t-1}}{\widehat{M}_{t-1}}$,

$$\Delta\log\widehat{M}_t = -\frac{\widehat{\Lambda}_{M,t-1}}{\widehat{\Lambda}_{t-1}}\frac{\widehat{\Lambda}_{t-1}}{\widehat{M}_{t-1}}\Delta\log\widehat{\Lambda}_{M,t}$$

I measure the ratio $\frac{\widehat{\Lambda}_{M,t-1}}{\widehat{\Lambda}_{t-1}}$ directly from the data and $\frac{\widehat{\Lambda}_{t-1}}{\widehat{M}_{t-1}}$ by computing the following integrals:

$$\frac{\widehat{\Lambda}_{t-1}}{\widehat{M}_{t-1}} = \frac{\int_{\hat{w}_{t-1}}^{\hat{h}_{t-1}}\left(\frac{1}{1-a}\right)^{\hat{\rho}}\left(\hat{h}_{t-1}-a\right)^{\hat{\rho}-1}\,\mathrm{d}a}{\int_{\hat{w}_{t-1}}^{\hat{h}_{t-1}} a\left(\frac{1}{1-a}\right)^{\hat{\rho}+1}\left(\hat{h}_{t-1}-a\right)^{\hat{\rho}-1}\,\mathrm{d}a}$$

Where $\hat{\rho} = \frac{\hat{\sigma}}{\hat{\sigma}-1}$.

B.2 Investment distribution in steady state

C Mathematical Appendix

C.1 Characterizing $G(z)$

The CDF of z is:

$$G(z) = 1 - \left(1 - \frac{1-\left(\frac{z_\ell}{z}\right)^{\frac{1}{\gamma}}}{1-\left(\frac{z_\ell}{z_h}\right)^{\frac{1}{\gamma}}}\right)^{\rho}$$

The CDF of the truncated Pareto is $F\left(z;\frac{1}{\gamma},z_{\ell},z_h\right)=\frac{1-\left(\frac{z_{\ell}}{z}\right)^{\frac{1}{\gamma}}}{1-\left(\frac{z_{\ell}}{z_h}\right)^{\frac{1}{\gamma}}}$, so:

$$G(z)=1-(1-F(z))^{\rho} \tag{48}$$

The survival function $\bar{G}(z)=1-G(z)$ is

$$\bar{G}(z)=\bar{F}(z)^{\rho}$$

The PDF of z is:

$$g(z)=\rho f(z)(1-F(z))^{\rho-1} \tag{49}$$

C.1.1 Moments and Hazard Rate of $G(z)$

With $\rho=\frac{\sigma}{1-\sigma}$, the mean of z is

$$\mathbb{E}_G(z)=H^{1-\sigma}L^{\sigma} \tag{50}$$

The hazard rate of G is

$$\begin{aligned}\frac{g(z)}{1-G(z)} &= \frac{\sigma}{1-\sigma}\frac{f(z)}{1-F(z)}\\ &= \sigma\frac{z_h^{1-\sigma}}{z\left(z_h^{1-\sigma}-z^{1-\sigma}\right)}\end{aligned}$$

$$\frac{g(z)}{1-G(z)}=\frac{\sigma}{z\left(1-\left(\frac{z}{z_h}\right)^{1-\sigma}\right)}$$

So the hazard rate of G is equal to a constant times the hazard rate of the truncated Pareto distribution with tail parameter $1-\sigma$. If $\sigma>\frac{1}{2}$ ($\sigma<\frac{1}{2}$), $G(z)$ has higher (lower) hazard rate than $F(z)$for all z, in which it has thinner (fatter) tails. For $\sigma=\frac{1}{2}$, the hazard rates (and hence distributions) are identical, so $G(z)$ is exactly the truncated Pareto for $\sigma=\frac{1}{2}$.