Firm Organization and Information Technology: Micro and Macro Implications*
Preliminary
Please do not distribute

Asier Mariscal

U. Carlos III-Madrid

February 15, 2018

The last 30 years in the US have seen: i- A rise in inequality with flat wages at the bottom of the wage distribution, ii-A reduction in the labor share (LS) of national income and, iii-A notable decline in the price of Information Technology (IT) capital. I show these facts can be rationalized with optimal capital choices of firms organized as hierarchies. The theory is consistent with capital-labor complementarity, for each skilled and unskilled labor, as in the data. In the model, cheaper IT capital enables faster managerial problem solving, and reallocates knowledge and wages away from the low levels of the hierarchy towards the higher ones. Larger firms use more IT capital, over time they reduce their marginal cost and expand more than smaller firms, hence capturing a larger share of aggregate value added. In the quantitative section, I discipline the model with firm-level moments and explain key features of firm organization. As firm value added increases there is 1-a lower LS of production costs, 2- a higher capital-labor ratio, and 3-higher wages conditional on the numbers of layers. Within a firm, as IT prices fall: 4- wage declines at low layers and increases at top layers, 5-increasing employment shares of managers; and, at the aggregate level, as a share of GDP: 6-increasing LS for managers, 7-declining LS for plant-level workers. Finally, the model explains: 8- the decline in the LS of value added of large firms, and 9-the decline in the aggregate LS. The last two results require both IT price declines and increasing mark-ups, the latter smaller than in recent demand-based explanations.

*I thank Pol Antras, Thomas Chaney, William Fuchs and Ben Moll for comments. The usual disclaimer applies.

1 Introduction

During the last 30 years there has been an intense penetration in the use of Information Technology capital, IT, strongly impacting the labor market. For example, for the US Autor et al. (2003) show that computers substitute unskilled labor and complements high-skilled labor. At the same time, inequality in the US is rising with the surprising feature that wages for the less skilled workers have fallen despite aggregate productivity growth, see Autor (2014). Changes in the distribution of income have also manifested in a lower share of national income for low-skill workers and a larger share for high-skill workers, with an overall decline in the labor share, IMF (2017). Many of these in the labor market trends are not unique to the US but shared by other European countries, see for example Autor (2010) and Karabarbounis and Neiman (2013).

This paper rationalizes these macroeconomic trends using a unified framework in which firms organize in hierarchies. In the model, the emergence of IT capital alters the relative price of problem solving at different layers of the firm, promoting the use of managers instead of floor level workers. More precisely, lower IT prices makes firms economize on knowledge at lower layers as it becomes cheaper to solve problems by managers. As a result, employment shares of managers and their wages increase, whereas those of lower layers fall, further generating an overall within-firm increase in capital-labor ratios. With lower costs and better technology, the largest firms expand their scale increasing their share of value added in the cross-section.

How much of the micro and macroeconomic trends in the labor market can be explained by firms' changing organization? In order to answer this question, I study the optimal organization of firms that choose labor, knowledge and capital as well as the number of layers in the hierarchy. Firms are heterogeneous in demand and respond to declining IT prices by optimally reorganizing. As a consequence, the macroeconomic changes are the result of within-firm choices and between firm reallocation, consistent with the micro-to-macro evidence in Oberfield and Raval (2014) and Kehrigy and Vincent (2014). The latter decompose the aggregate labor share decline between 1987-2007 in US manufacturing, and conclude that while reallocation of value added towards large firms plays an important role, these firms have also decreased their LS of value added (VA) substantially. Aligned with this evidence, my model will generate reallocation across-firms and within-firm changes in the LS.

This paper is related to four main strands of literature which I review next. The first

is the recent literature on the decline of the aggregate labor share of income. Karabarbounis and Neiman (2013) documented a decline in the labor share of income for a panel of countries and industries, a fact that has attracted widespread attention. Their explanation relies on an aggregate production function with a capital-labor elasticity larger than one, which generates enough substitution away from labor and towards capital as the relative price of the latter falls. However, their capital-labor elasticity is at odds with the empirical estimates in the literature, see for example Raval (2017a). Oberfield and Raval (2014) study the aggregate capital-labor substitution when firms are the relevant decision units. They show that within-industry heterogeneity matters for inference about the aggregate capital-labor substitution. They conclude that the bias of technical change is responsible for most of the change in the labor share as opposed to changes in relative factor prices. Like them, I use a micro-founded production function, which also features capital-labor complementarity. In contrast to them, in my model firms organize production in hierarchies choosing the amount of labor, knowledge and capital at each layer thereby deepening our understanding of how the organization of production has contributed to the change in the labor share in many countries. Koh et al. (2018) show that recent national accounting changes in how property products (IPP) capital are measured contribute substantially to the decline in the aggregate LS. My paper rationalizes the cross-sectionally heterogeneous shift to a more IPP capital-intensive economy.

DeLoecker and Eeckhout (2017) also focus on micro data to decompose the aggregate labor share into the underlying firm responses. Compustat data for the 1980-2014 period shows that firms' average mark-ups have increased, leading to a rise in profits and a decline in the factors' share of income. The increase in markups has occurred particularly in larger firms, a result consistent with the findings in Kehrigy and Vincent (2014). A related paper is Autor et al. (2017) who use micro data for a cross-section of countries and find: i-a tendency towards sales concentration at larger firms and ii-a robust correlation between sales concentration and the decline in the aggregate LS. While these papers provide important evidence, their demand-based explanations for the decline in the LS depend on their assumptions on the production side. In particular, each of them draws opposite conclusions as to how the elasticity of demand has changed in this period. I empirically discipline a theory of firm organization to make it consistent with an array of empirical findings and obtain a reduced role for demand in the decrease of the aggregate LS.

This paper also relates to the literature on hierarchies. The seminal paper is Robert E. Lucas (1978) who studies the role of the span of control of managers for the allocation of economic activity. He suggests an extension to further capture the complexity of how business firms are organized where, at each layer capital and labor are chosen. This paper does just that and examines it's quantitative implications in the light of recent economic trends. I build on a literature that was initiated by Garicano (2000), who models firms as an organization designed for optimal problem solving. Workers at lower layers solve the most common problems whereas managers only deal with the exceptions. My paper is closer to Caliendo and Rossi-Hansberg (2012), who study firm organization with international trade[1]. In the model, firms are heterogeneous in demand and choose the number of layers and, at each layer, knowledge and employees. While workers are homogenous ex-ante, ex-post they have different wages and knowledge. Importantly, Caliendo et al. (2015b) provide evidence based on French linked employer-employee data that is consistent with a large set of predictions in Caliendo and Rossi-Hansberg (2012)'s theory. They show that a firm-layer dummy explains a large share of the total variance in worker level hourly wages. This is important because it suggests that focusing on the hierarchies explains a large fraction of the cross-sectional determinants of wages. To the hierarchy literature I add optimal capital choices which have been absent from this strand[2]. The introduction of capital into the analysis is not only fundamental to understand how firms have changed their organization but also to understand the decline in the labor share.

A classic strand is macro in nature. Acemoglu and Restrepo (2018) study the race between automation and the creation of new tasks. They conclude that automation always reduces the aggregate LS and inequality increases during the transition through both channels. Their model can be viewed as a combination of task-based theories of the labor market with directed technological change models that allow capital accumulation[3]. The static part is similar to the model in Acemoglu and Autor (2011), which

[1] See also Antràs et al. (2006) and Antràs et al. (2008) study the implications of offshoring on wage inequality using a similar model of organizations. See also Marin and Verdier (2012) and Marin and Verdier (2014) for the connection between offshoring and decentralization of decisions in hierarchy models. See Fuchs and Garicano (2010) for an "in-house vs outsource" analysis in the context of hierarchies and Fuchs et al. (2015) for a study on optimal contracting in hierarchies with informational problems.

[2] A notable exception is Caliendo et al. (2015a), who estimate productivity using a short-run cost function, ie, with fixed capital

[3] On directed technical change models, see Acemoglu (1998), Acemoglu (2002), Acemoglu (2003), Acemoglu (2007), Kiley (1999), Caselli and Coleman (2006), Thoenig and Verdier (2003) among others.

provides a foundation for economic models that aim to explain a broad class of labor market outcomes[4]. Grossman et al. (2017) argue that that the productivity slowdown experienced by many countries in recent years may have caused the decline in labor's income. Their calibration suggests that a one percentage point decline in the productivity growth rate accounts for between half and all of the observed decline in the US labor share. Eden and Gaggl (2018) aggregate occupations into broad labor classes and emphasize the role of IT for the diverging routine and non-routine aggregate labor shares. Similarly, Barkai (2017) studies the factor shares of capital and labor separately and shows that the profit share has increased. The model in this paper is complementary to this strand of macroeconomics: it does not deal with aggregate growth, but in turn is able to explain a wide range of firm-level outcomes, including the cross-section of the LS of VA and the recent changes in how businesses organize their capital and labor.

I more specifically relate to a literature that studies the effects of IT adoption on firms and workers. In this line of work, Bloom et al. (2014) and Garicano and Rossi-Hansberg (2006) argue that there are two types of IT when it comes to their inequality consequences. First, IT that decreases knowledge access costs, like computers, and second IT that decreases communication costs. The first, computers, increase knowledge and wages at the floor level whereas the second, better communication, decreases them. In Garicano and Rossi-Hansberg (2006) firms are collections of heterogeneous workers who optimally choose to be a CEO or a worker. They show that their model responds to changes in the knowledge access cost and communication costs in a way that is consistent with the 1960 to 2000 aggregate wage inequality trends. The effect of the same primitives is the subject of Bloom et al. (2014), who model the consequences of such technologies in hierarchies, but like Garicano and Rossi-Hansberg (2006) not their adoption. In a management survey, they measure worker and manager autonomy as well as the span of control of managers and study their response to detailed technology use finding confirmation of their model. While I use a more coarse measure of IT, I study the optimal capital adoption and the consequences for several micro and macro trends, including the labor share.

Other papers have studied the effects of IT, without specifically focusing on hierar-

[4] There is an ample literature studying the effects of trade on inequality an issue with which I do not deal. See, for example,Feenstra and Hanson (1999), Spitz-Oener (2006), Goos et al. (2014), Grossman and Rossi-Hansberg (2008), Autor and Dorn (2013), Rodriguez-Clare and Ramondo (2010), Acemoglu et al. (2010), and many others.

chies. On the effects on skill, Autor et al. (2003) show that computers are substitutes for cognitive and manual routine tasks whereas it complements non-routine tasks, features that arise naturally in my model. Using linked employer-employee data from Norway, Akerman et al. (2015) exploit a quasi-natural experiment to show that broadband adoption in firms complements skilled workers in executing nonroutine abstract tasks, and substitutes for unskilled workers in performing routine tasks. Bresnahan et al. (2002) show that several firm-level measures of IT, like general purpose computing, the computerization of work or installed processing power are complementary to firm reorganization. They also show that such firm-level IT adoption measures are negatively associated with unskilled labor use. I obtain results in line with this literature and also quantitatively explain a rich set of firm level outcomes in the cross-section as well as aggregates.

This paper has 7 sections. On Section 2 I introduce the model, and it's solution and intuitions are described in Section 3 and 4. On Section, 5 I describe the empirical mapping to the data and on Section 6 I study the macro and micro trends that the model can explain using jointly demand elasticity and IT prices. Section 7 concludes.

2 Model

2.1 Preferences and workers

The economy is populated by N agents that inelastically supply one unit of labor. They are ex-ante identical, but acquire different knowledge ex-post, and with their wage they consume by maximizing utility

$$U(q(.)) = \left(\int_{\Omega} \alpha^{1/\rho} q(\alpha)^{\frac{\rho-1}{\rho}} g(\alpha) d\alpha \right)^{\frac{\rho}{\rho-1}} \tag{1}$$

where $q(\alpha)$ denotes the quantity of variety α consumed, $g(.)$ is the probability distribution over varieties in the set Ω, and $\rho > 1$ is the elasticity of substitution. Note that inverse demand is $p(\alpha) \propto q(\alpha)^{-1/\rho} \alpha^{1/\rho}$ and varieties with higher α deliver larger utility and hence, ceteris paribus, are demanded in larger quantities by agents.

In order to consume, agents use their available unit of time in the labor market, and obtain a wage. They can work in production or in the education sector, and in equilibrium they will earn w in both, regardless of the occupation they sort into. I

elaborate on this in the next subsection.

2.2 Firms

Entrepreneurs obtain a draw from $G(\alpha)$, which determines the level of demand for their product. Organizing production implies choosing optimally the number of managerial layers and the amount of each of three factors at each layer, namely, capital, k_l, labor, n_l, and knowledge, z_l. Labor and capital is used both in production at the plant and by managers solving problems.

The number of layers satisfies $1 \leq L \leq 3$, a range chosen following the empirical description of firm organization in Caliendo et al. (2015b). Within an organization, agents can be broadly grouped into workers and managers. The first are dedicated to production and the second to problem solving. Workers are combined with capital at layer 1, and the potential production input bundle is given by $y_1 \equiv \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{1/\beta_1}$. Per unit of input bundle one problem arises which if successfully dealt with, yields A produced quantity. Accordingly, total potential output is Ay_1.

Each production possibility is associated to a problem drawn from $F(x) = 1 - exp(-\lambda x)$. For a problem z drawn from that density to become realized output, the worker must have a knowledge set that includes such problem. Workers can deal with y_1 problems. Suppose workers learn the interval $[0, z_1]$, then $AF(z_1)y_1$ is produced with their knowledge and all problems above z_1 are sent to the managers at a layer above. Assume for exposition that such managers exist. The higher layers, $l > 1$, are specialized in problem solving and it's managers use IT capital to deal with and solve them. In layer 2 managers deal with unsolved problems from $l = 1$, and the required managerial input bundle $y_2 \equiv \left(k_2^{\beta_2} + n_2^{\beta_2}\right)^{1/\beta_2}$ is given by the following restriction,

$$\left(k_2^{\beta_2} + n_2^{\beta_2}\right)^{1/\beta_2} = \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{1/\beta_1} (1 - F(z_1)) \tag{2}$$

The left-hand side (LHS) of Equation 2 is the managerial input at $l = 2$, whereas the right-hand side (RHS) is the amount of problems unsolved at layer 1, which have been passed onto layer 2. To deal with such a volume of problems in layer 2 the firm optimally chooses the managers, their knowledge and IT capital. Assume managers learn interval $[z_1, z_1 + z_2]$, the amount of problems passed onto layer 3 are $y_1(1 - F(z_2))$, i.e. the unsolved problems at all previous layers; those above z_2. More generally, at layer $l > 1$,

managers learn $[Z_{l-1}, Z_l]$ with $Z_l = \sum_{k=1}^{l} z_k$, and satisfy $\left(k_l^{\beta_l} + n_l^{\beta_l}\right)^{\frac{1}{\beta_l}} = y_1(1 - F(Z_{l-1}))$ with $\beta_l \equiv \frac{\sigma_l - 1}{\sigma_l}$. The process continues until it reaches the top layer, the CEO. As the process of adding layers and knowledge continues more of the potential production is realized, such that for a firm with total knowledge Z_L, realized production is $AF(Z_L)y_1$. Section 3 describes the problem and solution of an arbitrary firm with it's full set of constraints.

So far I have introduced the benefits but knowledge also entails a cost, like capital and labor. The cost of acquiring knowledge z is assumed to be wcz. The cost of knowledge is proportional to w because learning requires teachers in the schooling sector who are paid w. With a decreasing density $F(.)$, economizing on knowledge implies that workers at lower layers learn the more common problems, whereas those at higher layers learn the more exceptional problems. Agents acquire knowledge that the firm will exactly compensate leaving them indifferent across sector and occupations with a net wage of w.

3 Model Solution

The owner of the firm solves the organization problem in two steps. First, conditional on L, he chooses production factors, capital, labor and knowledge at each layer. Second, he compares profits across organizational modes, i.e. across the different L=1,..,4. On a first step, the problem of an entrepreneur is to solve, for fixed L:

$$\max_{q,\{z_l,n_l,k_l\}_{l=1}^{L}} p(q)q - \sum_{l=1}^{L} \left(n_l w(cz_l + 1) + p_l k_l\right) \tag{3}$$

subject to:

$$\begin{cases} p(q) = \left(\frac{q}{\alpha R}\right)^{-1/\rho} & \\ q = A[1 - \exp(-\lambda Z_L)]\left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{\frac{1}{\beta_1}} & : \phi \\ \left(k_l^{\beta_l} + n_l^{\beta_l}\right)^{\frac{1}{\beta_l}} = \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{\frac{1}{\beta_1}} \exp(-\lambda Z_{l-1}), L > l > 1 & : \psi_l \\ k_L^{\beta_L} = \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{\frac{1}{\beta_1}} \exp(-\lambda Z_{l-1}), L > l > 1 & : \psi_L \\ n_L = 1 & \\ z_l > 0, \forall l \geq 1 & : \theta_l \end{cases} \tag{4}$$

Profits are the objective function, where the first term is the revenue and the second is total cost. The latter is composed of the wage bill and capital spending, where p_l are prices of capital at each layer. The first restriction is the demand curve of the firm for it's α variety. The second is the production function. The rest are the managerial input constraints.

The restriction at the CEO level is different as it assumes CEO time is fixed, unlike managerial labor in previous layers, which implies there is no capital-labor substitution; Section 4.3 discusses the modeling of capital and labor at the top, which is chosen in order to match empirical moments but previewing the discussion in that Section is useful. Fixing CEO time will make wages grow with firm size conditional on L, whereas allowing capital at the top allows CEOs to increase their span and the average cost increases less steeply with firm size. That IT capital increases managerial span of control should not be controversial, particularly when interpreted as all the information systems in the corporation that relieve the CEO from time consuming activities. It is likely that the recent wave of IT capital allow for CEOs to focus on what really is exceptional problems and leave the rest to automated systems and lower managerial layers.

In what follows, in an effort to further discipline the quantitative results, I assume production workers use production capital, k_1 and managers at different layers use potentially different amounts of IT capital k_l but with common price p_2. Moreover, $\beta_l = \beta_2$ for $L > l > 1$ in the quantitative section.

In order to understand the solution, it is instructive to solve the cost minimization problem first. By obtaining the first order conditions (FOCs) and combining them one can gain relevant intuitions on the trade-offs faced in organizing production. For example, except at the top layer, the capital-labor trade-off is solved at every layer:

$$\frac{k_l}{n_l} = \left(\frac{p_l}{w(cz_l + 1)}\right)^{-\sigma_l} \tag{5}$$

This is the familiar condition arising in two-factor CES production functions with elasticity parameter σ_l with one difference: in this model, knowledge is chosen optimally and hence wages are not exogenous to the firm. The top layer is different because CEO time is fixed:

$$k_L^{\beta_L} = \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{\frac{1}{\beta_1}} \exp(-\lambda Z_{l-1}) \tag{6}$$

Intuitively, since labor is not variable at the top, k_L is a "residual", i.e. it does not comes from a FOC but from the time constraint of the CEO. The effect of IT capital at the CEO level is to relieve the time constraint he faces.

Since a defining element in this theory is knowledge, it is useful to gain intuition of its optimal choice. From FOCs, we obtain that the knowledge trade-off across layers 1 and 2 is:

$$wcn_1 - \psi_2 \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{\frac{1-\beta_1}{\beta_1}} \lambda \exp(-\lambda z_1) = wcn_2 \tag{7}$$

where ψ_l is the multiplier associated with layer l "time" or inputs, and can be interpreted as the cost per extra unit of managerial product at l. Mathematically,

$$\psi_l = \begin{cases} \dfrac{w(cz_l + 1)}{\left(k_l^{\beta_l} + n_l^{\beta_l}\right)^{\frac{1-\beta_l}{\beta_l}} n_l^{\frac{-1}{\sigma_l}}} \\ \dfrac{p_l}{\left(k_l^{\beta_l} + n_l^{\beta_l}\right)^{\frac{1-\beta_l}{\beta_l}} k_l^{\frac{-1}{\sigma_l}}} \end{cases} \tag{8}$$

where $\beta_l \equiv \frac{\sigma_l - 1}{\sigma_l}$ for $l > L$. With this in mind, it is apparent that equation 7 shows the trade-off in knowledge across layers: the marginal cost of knowledge in layer 1 net of the reductions in communication costs in layer 2 (LHS) equals the marginal cost of knowledge in layer 2 (RHS). Note that since the multiplier is associated to both capital and labor, there are two channels through which this trade-off operates. In particular, the same equation where $\psi_2 > 0$ is substituted on the "capital channel" becomes,

$$wcn_1 - \frac{p_2}{\left(k_2^{\beta_2} + n_2^{\beta_2}\right)^{\frac{1-\beta_2}{\beta_2}} k_2^{\frac{-1}{\sigma_2}}} \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{\frac{1}{\beta_1}} \lambda \exp(-\lambda z_1) = wcn_2 \tag{9}$$

This equation illuminates an important insight: A reduction in IT prices p_2, makes layer 1 input relatively cheaper, and tends to increase knowledge at layer 1 and decrease it at layer 0. This forms the basis of the mechanism that generates inequality within the firm.

It is useful to further operate on the FOCs and restrictions of the cost minimization problem to obtain the conditional factor demands,

$$
\begin{cases}
n_1 = \dfrac{(w(cz_1+1))^{-\sigma_1}}{A[1-\exp(-\lambda Z_L)]} P_1^{\sigma_1} q \\
k_1 = \dfrac{p_1^{-\sigma_1}}{A[1-\exp(-\lambda Z_L)]} P_1^{\sigma_1} q \\
... \\
n_l = \dfrac{(w(cz_l+1))^{-\sigma_l}}{A[1-\exp(-\lambda Z_L)]} P_l^{\sigma_l} \exp(-\lambda Z_{l-1}) q \\
k_l = \dfrac{p_l^{-\sigma_l}}{A[1-\exp(-\lambda Z_L)]} P_l^{\sigma_l} \exp(-\lambda Z_{l-1}) q \\
... \\
k_L^{\beta_L} = \dfrac{\exp(-\lambda Z_{L-1})}{A[1-\exp(-\lambda Z_L)]} q
\end{cases}
\tag{10}
$$

where $P_l = \left(p_l^{1-\sigma_l} + (w(cz_l+1))^{1-\sigma_l}\right)^{1/(1-\sigma_l)}$

Conditional factor demands in Equation 10 are not unfamiliar since they inherit the CES structure to a large extent. Knowledge, which will be optimally chosen in this theory, is the differing element. It appears in three places: wages, the denominator and as a shifter in the numerator, inside the exponential term. The reason for it's appearance in wages is clear and the intuition for the latter two is also simple. Knowledge acts as a Hicks-neutral TFP term in that more of it delivers more production, similar to the role of A. On the other hand, the numerator contains the intuition for the trade-off between knowledge vs labor and capital. The more knowledge at layers lower than k, the less factors at layers above k are needed to attain production q.

The maximization problem of the firm given L can be solved subject to the conditional factor demands in Equation 10 and delivers the quantity of production, as well as, knowledge, labor and capital at each layer.

4 Understanding the Mechanisms

4.1 Firm Organization Choice

The previous Section provided intuition for the solution for fixed L, and I now turn to illustrating the optimal choice of L. In the model, entrepreneurs choose among 3 organizational modes; I exclude self-employment for now. I choose the range for L

following the detailed empirical mapping of occupations to layers in business hierarchies in Caliendo et al. (2015b); more on this in Section 5.1.

Firms choose $2 \leq L \leq 4$, and they do so by maximizing profits across L. Since profits are intimately related to costs, in this Section I graphically illustrate cost and profits for $L = 2, 3$ with optimal $q(\alpha)$. Before turning to the figures, the cost function $C(q; z, L)$ for arbitrary L is given by Equation 11. For what follows, it will be useful to note that the last term is related to CEO time and is convex on q since I will assume $\beta_L > 1$.

$$C(q; z, L) = \sum_{l=1}^{L} (w_l n_l + p_l k_l) =$$
$$\sum_{l=1}^{L-1} \left(\frac{(w(cz_l + 1))^{1-\sigma_l}}{A[1 - \exp(-\lambda Z_L)]} P_l^{\sigma_l} \exp(-\lambda Z_{l-1}) + \frac{p_l^{1-\sigma_l}}{A[1 - \exp(-\lambda Z_L)]} P_l^{\sigma_l} \exp(-\lambda Z_{l-1}) \right) q + \cdots$$
$$+ \left(\frac{\exp(-\lambda Z_{L-1})}{A[1 - \exp(-\lambda Z_L)]} q \right)^{\frac{1}{\beta_L}} \tag{11}$$

Figure 1 shows average cost for given L have a U-shape, due to CEO time being fixed. Because his time is fixed, as α increases, the firm wants to produce more, but it has to do so by using any factor except CEO time. With such fixed factor, the marginal cost increases with the quantity produced and so does the average cost; this is illustrated on the RHS of Figure 1, on higher α values. On the other hand, even if knowledge was set to zero, there is still a wage cost of w per worker. This can be thought of as a "fixed" cost so that, naturally, on low values of α as production increases average cost falls. Figure 2 shows how increasing marginal costs make average cost U-shaped.

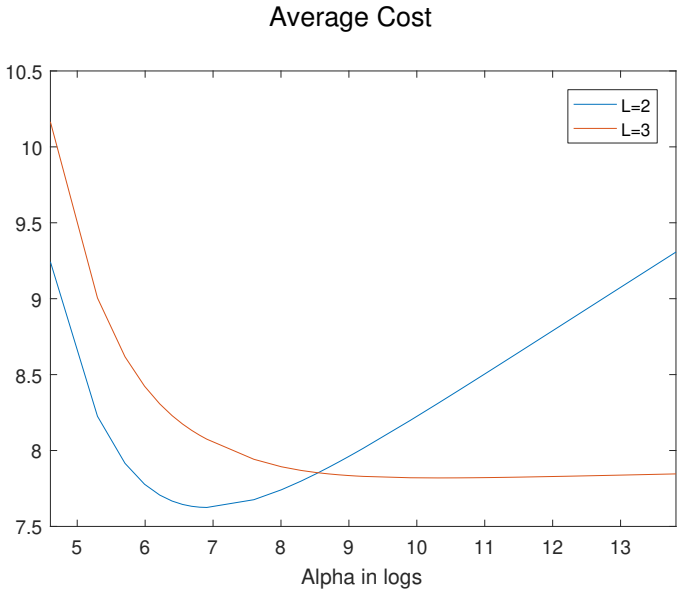


Figure 1: Average costs in the α cross-section conditional on $L = 2, 3$.

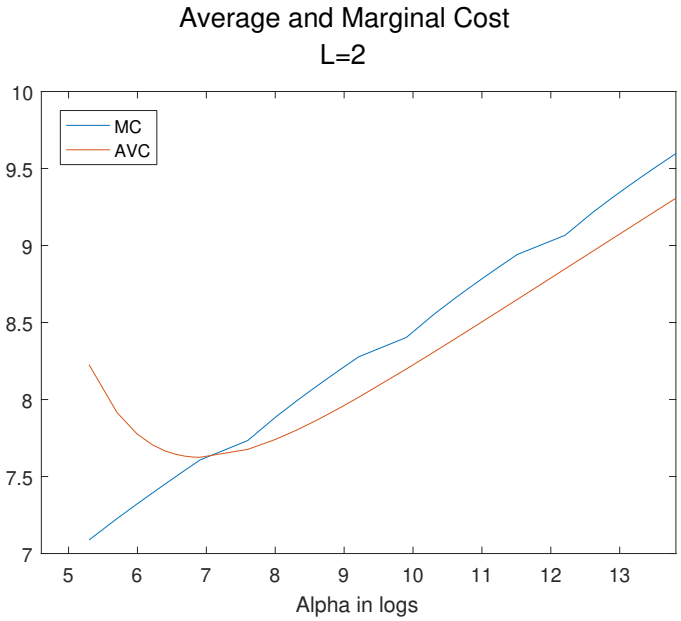


Figure 2: Average and marginal costs conditional on $L = 2$.

With the determinants of cost in mind, thinking about the optimal L comes naturally. Figure 3 illustrates profits across α. Note that smaller α firms prefer simpler organizations, with $L = 2$, and larger α entrepreneurs choose $L = 3$.

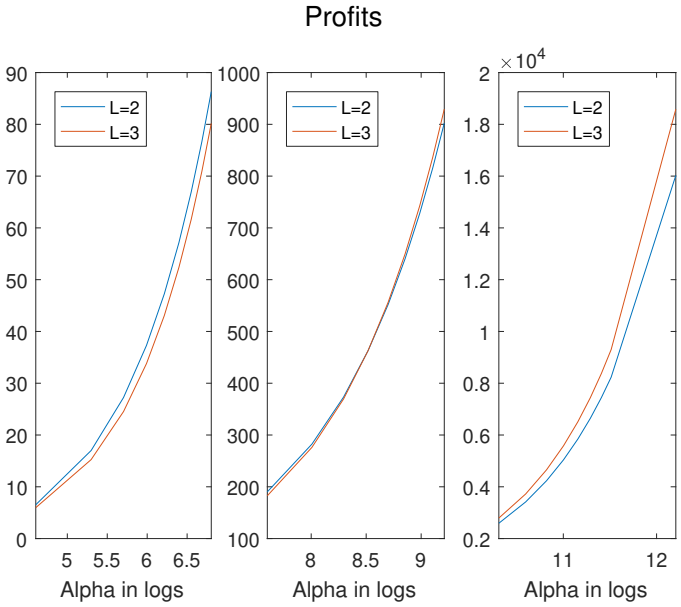


Figure 3: Profits in the α cross-section conditional on $L = 2, 3$.

Figure 4 concludes the section noting the effects of increasing idiosyncratic demand α conditional on $L = 3$. Both quantities and revenues increase, since despite the increasing marginal cost, and hence price, consumers derive larger utility from higher α varieties. To satisfy the larger demand, more factors are used, including knowledge, and wages at all layers increase; and as wages increase, the firm increases substitutes towards capital increasing it's capital-labor ratio. All of these correlations are observed in the data and this theory provides a rationale.

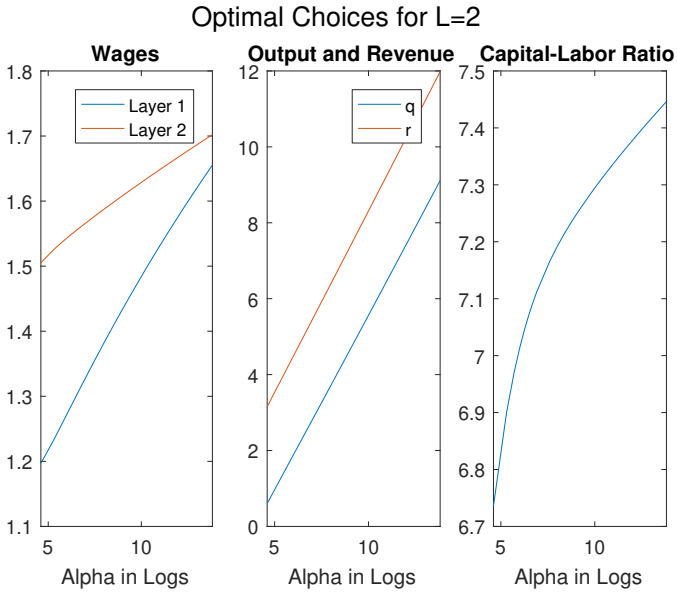


Figure 4: Optimal choices in the α cross-section conditional on $L = 2$. Vertical axis in logs.

4.2 The Effects of Declining IT Prices

In this Section I describe the within-firm and across-layers response to declining IT capital prices for the median firm with optimal $L = 3$. Several authors have documented the IT capital price evolution, for example recently Eden and Gaggl (2018). In their paper, they report a decline of roughly 2/3-ths in the price of IT capital relative to that of final goods. Figure 5 shows the response of wages at the three layers. All figures in this section report changes relative to the baseline year, with relatively high IT capital prices. As suggested by Equation 7, a decline in IT prices triggers an optimal reallocation of knowledge across layers: managerial problem solving becomes cheaper, hence they learn a larger set of problems, some of which used to be solved by the plant-level workers. Firms find it cheaper to allocate knowledge to few managers instead of many workers, because now the former can leverage their time with IT capital.

With lower IT capital prices, the firm has a lower marginal cost and expands demanding more labor and capital. Regarding labor, only the lowest and intermediate layers respond, since the CEO time is fixed. On Figure 6 intermediate managers increase their employment share whereas blue-collar workers decrease it. Since labor is complementary to capital, lower IT capital prices induce more demand of k_2 which

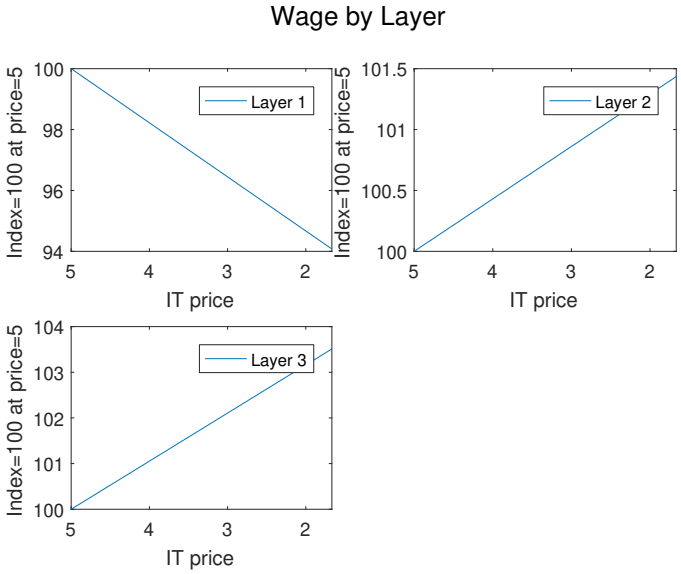


Figure 5: Wage response to IT capital price decline. Relative to baseline price.

in turn increases the demand for managers, n_2. Another way of thinking about it is through the conditional factor demands, Equation 10. There are two effects. First, both n_1 and n_2 increase as the firm optimally increases it's scale. Second, since knowledge at lower layers has gone down, more inputs at higher layers are used. Overall, the increase in managerial labor is larger than that of plant workers and hence the firm-level employment shares move in opposite directions.

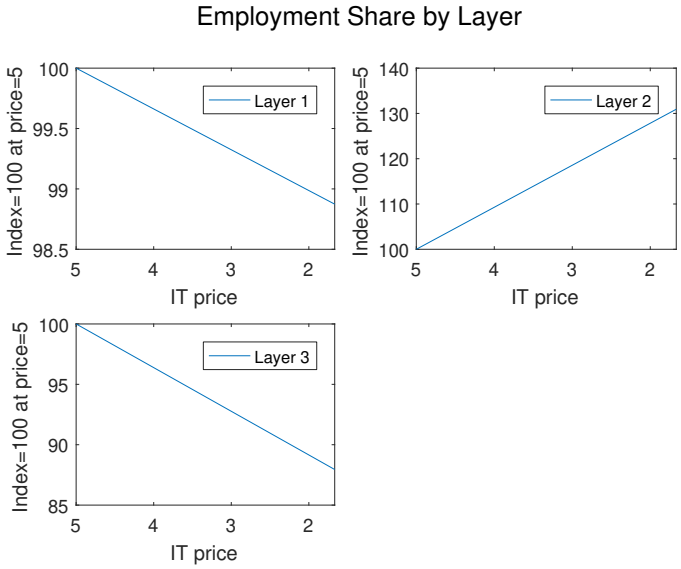


Figure 6: Employment share response to IT capital price decline. Relative to baseline price.

Capital-labor ratios respond as shown in Figure 7. Since wages increase at layer 2 and decrease at layer 1, optimal substitution makes the capital-labor ratios to move in opposite direction in each of those layers. As the firm expands, the firm responds allocating more capital to the CEO as well.

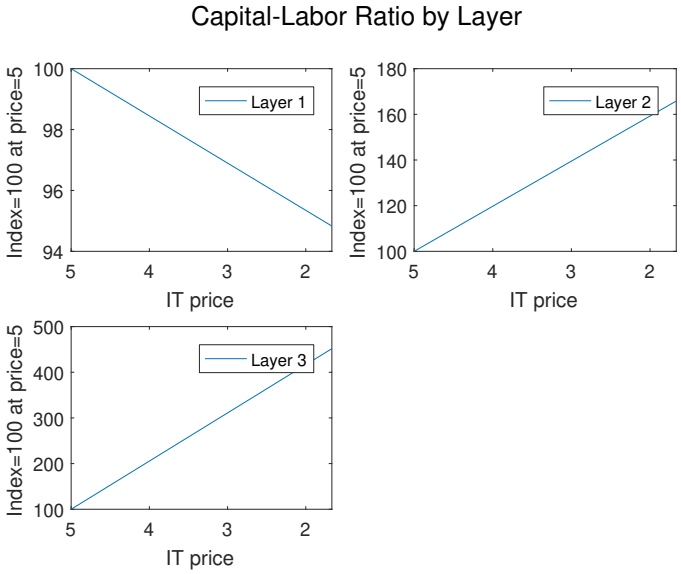


Figure 7: Capital-labor ratio response to IT capital price decline. Relative to baseline price.

With the described employment and wage responses, naturally the share of plant-level workers in total wage bill falls whereas that of managers increases. But how about the firm-level labor share of production costs? In order to determine how the LS of cost responds to falling IT capital prices, note:

$$\frac{dLScost}{dp_2} = \frac{1}{TC}\left[\left(\underbrace{\sum_{l=1}^{L}\frac{dw_l}{dp_2}n_l}_{(1)} + \underbrace{\sum_{l=1}^{L}\frac{dn_l}{dp_2}w_l}_{(2)}\right)(1-LSC) - LSC\left(\underbrace{\sum_{l>1}^{L}k_l}_{(3)} + \underbrace{\sum_{l=1}^{L}\frac{dk_l}{dp_2}p_l}_{(4)}\right)\right] \tag{12}$$

where $LSC = \sum_{l=1}^{L}\frac{w_l n_l}{\sum_{j=1}^{L}(w_j n_j + p_j k_j)}$, and $TC = \sum_{l=1}^{L}(w_l n_l + p_l k_l)$.

Equation 12 highlights the conceptually different channels with (1) being one that differentiates this model from a more standard CES production function: optimal knowledge. While all terms are influenced by optimal knowledge only the underscored one is unique to the theory in this paper. Note also that (4) also includes the reaction of capital at the top layer, a margin that also plays an important role. From this

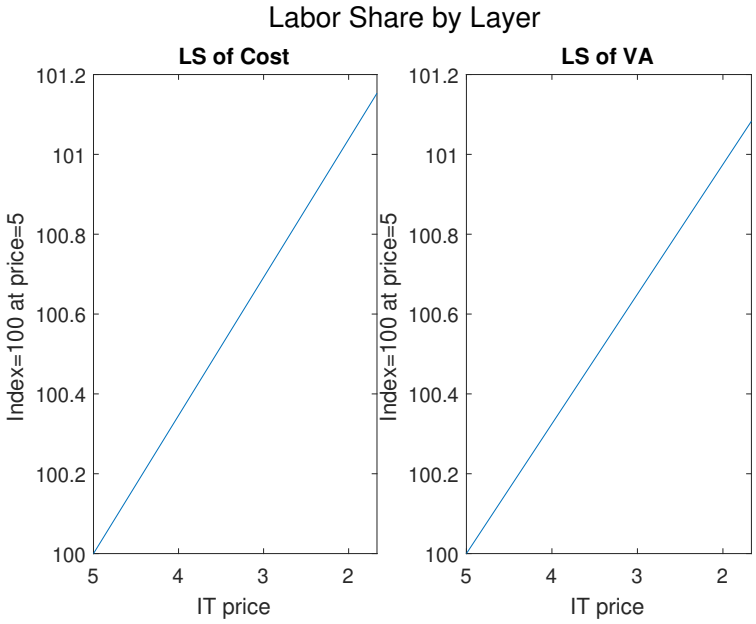


Figure 8: Labor share response to IT capital price decline. Relative to baseline price.

equation, we gain several pieces of intuition. First, as we have seen above, knowledge and wages respond optimally to capital prices and move in opposite directions at $l = 1$ versus $l = 2, 3$. However, as long as there are more workers at the lower layer than at higher levels, term (1) will tend to be positive. Hence, the response of z contributes to a decline in the LSC when p_2 falls.

Terms (2) and (4) capture the response of conditional factor demands to changing prices. Overall, in the calibration, they are negative due to the increased firm production. Finally, (3) is positive. Overall, each of the two summands (1+2) and (3+4) feature conflicting effects and are influenced by the level of the LSC. When the LSC is large, (3+4) play a larger role, and when LSC is low, as it happens for large α firms, (1+2) matters more.

$\frac{dLog(LScost)}{dLog(p_2)}$	**Total Change=**	**(1)**	**(2)**	**(3)**	**(4)**
0.50	100	1.71	-2.92	-3.67	104.87

Table 1: Decomposition of the change in the labor share to 1% lower IT capital price. Each column reports values as percentage of the total change and add up to 100. Reported values are for the median α firm for $L = 2$.

It is instructive to decompose the change in the labor share of cost into each of it's constituting elements as reported in Equation 12. For simplicity I focus on the median α firm for $L = 2$. Table 1 reports the elasticity of the LSC to the IT price and the contribution of each channel to the change in the labor share of cost as a result of decline in IT capital prices of 1%, all in percentage terms. As reported on the first column, the LSC decreases 0.5% as IT prices fall 1%. In the remaining columns, negative values on the table contribute to an increase of LSC and positive values have the opposite interpretation and they all add up to the total change, normalized to a 100. Notice how (1) and (4) contribute substantially to a decline in the LSC dominating (2) + (3) which contribute to an increased LSC.

4.2.1 The Heterogeneous Effect of IT Prices across Firm Organizations

Table 2 compares the response of firms with organizations $L = 2$ and $L = 4$. I use the calibrated parameters and focus on the median firm with each organization. In both cases, wages at the lowest layer fall, though substantially more in the $L = 4$ firm. The large decline in wages at the bottom of the wage distribution in large firms is consistent with the results reported in Song et al. (2015). On the other hand, wages increase for both the $L = 2$ CEO and the last two layers in $L = 4$. The large firm expands quantities and revenues by more than the L=2 firm, as in the documented increased sales concentration at larger firms in Autor et al. (2017). Through purely a decline in IT prices the LS of VA and LSC at large firms increases, a feature which is at odds with the results in Kehrigy and Vincent (2014), who document a substantial reduction in the LS of VA in the largest firms. This fact will allow me to identify changes in the

elasticity of demand in Section 5.2 when calibrating the model.

Organization:		
Variable	**L=2**	**L=4**
Wage layer 1	-3.75	-6.29
Wage layer 2	2.75	-2.6
Wage layer 3	.	3.4
Wage layer 4	.	4.37
Quantities	6.08	10.27
Revenues	3.83	6.46
LS Cost	-0.26	1.15
LS of VA	-0.46	1.15

Table 2: Effects of declining IT capital price across firms with different optimal organizations. Responses of calibrated model for IT price changes from 1980 to today. Values are changes in percentage for the median α firm for each organization.

4.3 The Relation Between Firm Scale and the Micro Labor Share

Oberfield and Raval (2014) emphasize that understanding the micro labor share in the cross-section of firms is crucial to understand the macro labor labor share. In this subsection I discuss the mechanisms that determine the LSC in my model and also compare it to two other natural benchmarks, an exercise that allows to perform model selection. In this model a firms' labor share response to α is,

$$\frac{dLScost}{d\alpha} = \frac{1}{TC}\left[\sum_{l=1}^{L}\left(\underbrace{\frac{dw_l}{d\alpha}n_l}_{(1)} + \underbrace{\frac{dn_l}{d\alpha}w_l}_{(2)}\right)(1-LSC) - LSC\underbrace{\sum_{l=1}^{L}\frac{dk_l}{d\alpha}p_l}_{(3)}\right] \tag{13}$$

The three terms are positive. Again, (1) is specific to the theory in this paper as it captures the response of knowledge to scale: as α increases, firms respond by increasing knowledge so they can produce more. Hence, labor becomes more expensive and firms

switch away from labor and towards capital. These mechanisms highlight a unique feature of this model: as a consequence of the knowledge margin, $\frac{dLScost}{d\alpha}$ can be seen as the combined effects of wages and scale.

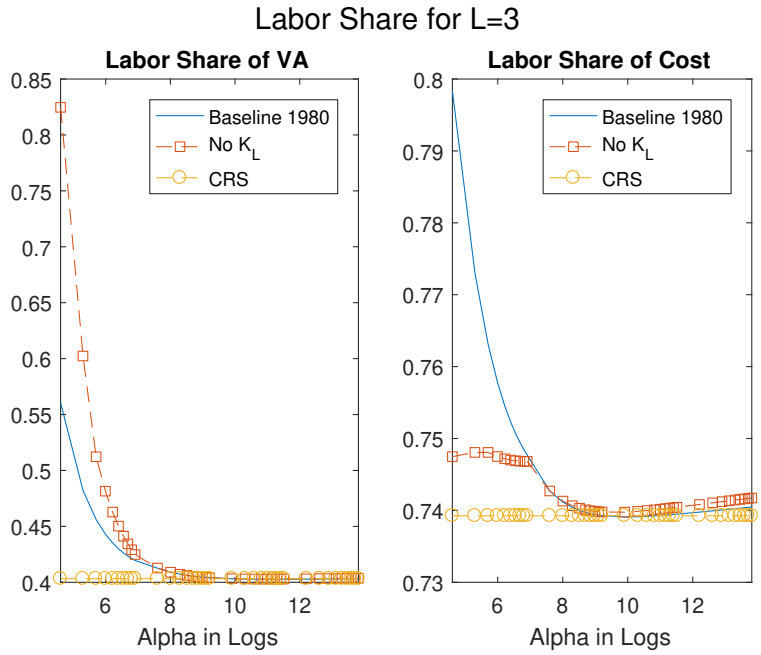


Figure 9: Labor share of cost and the α cross-section. The figure compares the benchmark model to two alternatives: i-a model with constant returns to scale at all layers (CRS), and ii-a model without capital at the CEO level (NOk_L). Same calibration except at top layer.

It is useful to compare the LSC in the benchmark model and to two useful alternatives: i-a model with constant returns to scale at all layers (CRS), and ii-a model without capital at the CEO level (NOk_L). Constraints at the top layer for each case are given by:

$$\begin{cases} \left(k_L^{\beta_L} + n_L^{\beta_L}\right)^{\frac{1}{\beta_L}} = \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{\frac{1}{\beta_1}} \exp(-\lambda Z_{L-1}) & (CRS) \\ 1 = \left(k_1^{\beta_1} + n_1^{\beta_1}\right)^{\frac{1}{\beta_1}} \exp(-\lambda Z_{L-1}) & (NOk_l) \end{cases} \tag{14}$$

The LSC for each model is shown in Figure 9, where the three models share the calibration except at the top layer constraint. Let's focus first on model (CRS). It is easy to see that it delivers a constant LSC across α, once we note that the cost function, before z is optimized, is independent of scale:

$$C(q;z,L)=\sum_{l=1}^{L}(w_l n_l+p_l k_l)=$$
$$\sum_{l=1}^{L}\left(\frac{(w(cz_l+1))^{1-\sigma_l}}{A[1-\exp(-\lambda Z_L)]}P_l^{\sigma_l}\exp(-\lambda Z_{l-1})+\frac{p_l^{1-\sigma_l}}{A[1-\exp(-\lambda Z_L)]}P_l^{\sigma_l}\exp(-\lambda Z_{l-1})\right)q \tag{15}$$

As a consequence, z is chosen irrespective of α or the production scale. Wages do not change with α nor does the LSC, features that are at odds with the empirical literature.

On the other hand, under model (NOk_L) firms cannot respond to a larger α by increasing capital at the top, hence they miss a part of margin (3) in Equation 13, i.e. $\frac{dk_L}{d\alpha}p_2$. Since (3) contributes to a negative overall sign, the net $\frac{dLScost}{d\alpha}$ still falls as α decreases but much less so.

In summary, the span of control of the CEO is a key identifying parameter. A final lesson from the LHS of Figure 9 is that solely focusing on LS of VA can produce misleading inference on the drivers of the aggregate LS.

4.4 Equilibrium

The equilibrium in this economy is a set of prices (w, p_l, P) such that:

- Consumers maximize utility subject to their budget constraint, as described in Section 2.1.
- Firms maximize profits, as described in Section 2.2.
- Labor, capital and goods markets clear.

For now in the quantitative Section 5 I will focus on Partial Equilibrium results.

5 Quantitative Analysis

5.1 Empirical Mapping of Occupations to Hierarchy Layers

The mapping of occupations to layers, follows Caliendo et al. (2015b) and allows me to have an empirical mapping from the theory in this paper to the data, and to focus on $2 \leq L \leq 4$. They separate workers according to their hierarchical level in the organization, that is, on the basis of the number of layers of subordinates that they have below them. In their French manufacturing data, the occupational classification is named PCS-ESE and includes five occupational categories as presented in Table 3.

Category	Description
2.	Firm owners receiving a wage, which includes the CEO or firm directors;
3.	senior staff or top management positions, which includes chief financial officers, heads of human resources, and logistics and purchasing managers;
4.	employees at the super visor level, which includes quality control technicians, technical , accounting, and sales supervisors;
5.	qualified and nonqualified clerical employees, secretaries, human resources or accounting employees, telephone operators, and sales employees;
6.	blue-collar qualified and non-qualified workers, welders, assemblers, machine operators, and maintenance workers.

Table 3: Occupations and Hierarchical Layers.

Through out the paper they merge classes 5 and 6 since the distribution of wages of workers in these two classes is remarkably similar, indicating similar levels of knowledge.

5.2 Calibration

In Table 4 there are two sets of parameters according to whether they are are calibrated externally or internally. Several parameter values are normalizations follow Caliendo and Rossi-Hansberg (2012), namely (A, R, P). At this stage, in partial equilibrium, their role is similar and related to average firm scale. The capital-labor elasticities come from Raval (2017b), and are aligned with economists' prior that both skilled and unskilled labor are complementary to capital though the former more so than the latter,

Autor (2015). I borrow the change in IT capital prices from Eden and Gaggl (2018) who use BEA data.

There is a set of parameters that are chosen to match data moments. The relative complexity of the model makes it difficult to identify each parameter separately, but I will provide intuition for how each parameter is informative of each simulated moment. ρ, the demand elasticity shifts up firms' LS of VA, by altering their mark-ups. I choose ρ to target values at the top of the size distribution and match the values reported in Kehrigy and Vincent (2014) for what they call "Highly-productive plants"; the model equivalent of the latter are $L = 4$ firms. In the calibration, I assume in 1980 mark-ups were constant in the cross-section at the initial value, whereas for 2015 I allow the value of ρ to gradually decrease with α starting at the same 1980's value for the smallest firm and evenly lowering it as α increases. The largest firm has ρ given by the second value on brackets[5].

Increases in w, the wage, makes the LS increase due to the capital-skill complementarity so I choose it to match a cross-sectional average of 0.5. As shown in Section 4.3, β_L generates dispersion in the LS of cost, so I choose it to match the slope of a regression of the factor cost ratio $FC \equiv \frac{CapitalCost}{WageBill}$ on value added, which I obtain from Raval (2017b). Capital prices also play a role in the degree of LSC dispersion delivered by the model in the cross-section. Higher prices makes substitution away from capital less profitable and, hence, there is less dispersion. I use the $p75/p25$ values from the Factor Cost Ratio in Raval (2017b). c and λ play similar roles as they both govern the importance of knowledge. c is the cost per unit of knowledge and reducing it makes knowledge less relevant which favors simpler organizations. λ is a shifter of the density of problems, with larger values implying more mass at low z values, reducing the incentives to add layers. Finally, the α distribution is assumed to be log-normal with parameters (μ_α, std_α)). These four parameters are chosen to match the US aggregate employment shares by firm size-class obtained from the BLS as well as the employment of the smallest $L = 2$ firm.

[5]See Lashkari and Mestieri (2017) for preserving nice aggregation properties while allowing for heterogeneous mark-ups.

Parameter	Value	Description	Source/Target
		Calibrated Externally	
A	5	TFP	Normalization
$p_l = p_2$, $l > 1$	[5, 5/3]	Capital price at l=2,3,4. Change.	Eden and Gaggl, (2017)
σ_1	0.87	Capital-labor elasticity at plant	Raval, (2011)
σ_l, $l > 1$	0.45	Capital-labor elasticity at l, $L > l > 1$.	Raval, (2011)
P	1	Price Index	Normalization
R	8.5	Aggregate income	Normalization
		Calibrated Internally	
ρ	[2.8, 2.2]	Demand elasticity	Average LS of VA at L=4
w	2	Wage	LS of cost, average
β_L	0.3	Capital exponent at layer L	Slope(log(FC),log(VA))
p_1	0.01	Plant-level capital price	p75/p25 FC distribution
c	0.5	Training cost	Employment at smallest $L = 2$
λ	4	Mean of $F(.)$	Employment share by firm size (5 bins)
μ_α	6.5	Mean of α	Employment share by firm size (5 bins)
$std(\alpha)$	2.2	Standard deviation of α	Employment share by firm size (5 bins)

Table 4: Full set of calibrated parameters in the baseline.

Table 5 shows the targeted moment data and simulated values, and their source[6]. The simulated and real data moments are close in this preliminary calibration. There are several margins with room for improvement. First, the average LS of VA at top firms in 1980 needs to be slightly higher and the value in 2015 lower. The correlation between FC and value added is relatively close though the model falls a bit short on generating FC dispersion measured as $p75/p25$. The simulated employment share distribution has a large explained variance.

[6] * indicates the moment is under construction.

Moment	Model	Data	Data Source
Firm-level moments			
Average LS of VA at L=4	[0.42, 0.38]	[0.52, 0.3]	Kehrigy and Vincent, (2017)
LS of cost, Average	0.67	$*$	Pending
Slope(log(FC),log(VA))	0.16	[0.05-0.1]	Raval, (2017)
p75/p25 FC distribution	1.44	2.1	Raval, (2017)
Employment share by firm size (5 bins)	R^2= 0.83		BLS

Table 5: Targeted Moments: Data and simulated. Data on levels are from 1980 and data on changes use the 1980-2015 period; Alternatively I use the most proximate years available in the cited papers.

While in my current calibration the demand elasticity in 1980 is 26% lower than the more standard value of 3.8, my implied increase in aggregate mark-ups during the period is 10%. On the other hand, for example DeLoecker and Eeckhout (2017) obtain a 41%. Note however that my results are preliminary[7].

6 Aggregate and Firm-level Changes from 1980 to Today

This section describes the results for the calibrated economy. It views the changes from 1980 to today as both decreasing IT capital prices and decreasing elasticity of demand, the latter more so at the larger firms. Figure 10 shows how heterogeneous firms optimally choose L. Firms with larger α sell more, and since their scale is large, they optimally add layers, which increases the fixed part w but delivers lower average costs. Lower IT capital prices makes it cheaper to increase layers, as firms increase capital-labor ratios.

[7] Calculations come from $(1.72 - 1.56)/1.56 = 0.1$ and $(1.67 - 1.18)/1.18 = 0.41$.

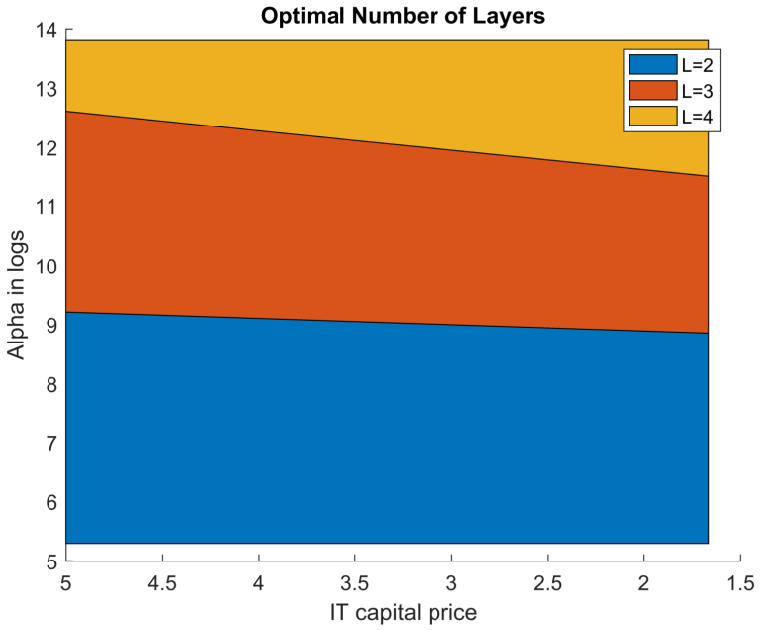


Figure 10: Optimal number of layers as function of (α, p_2).

Figure 11 shows the distribution of the LS of VA and LSC as function of α. In the cross-section of firms, as value added increases the LS decreases. This is the result of the mechanisms described in Section 4.3 and self-selection into L^*, which causes minor discrete jumps. The figure shows the distribution in 1980 and today. As explained in Section 4.2, the LSC tends to increase as p_2 falls, due to the capital-labor complementarity present in the model. On the right-hand figure, because mark-ups are increasing with α the distribution of LS of VA shifts downward gradually.

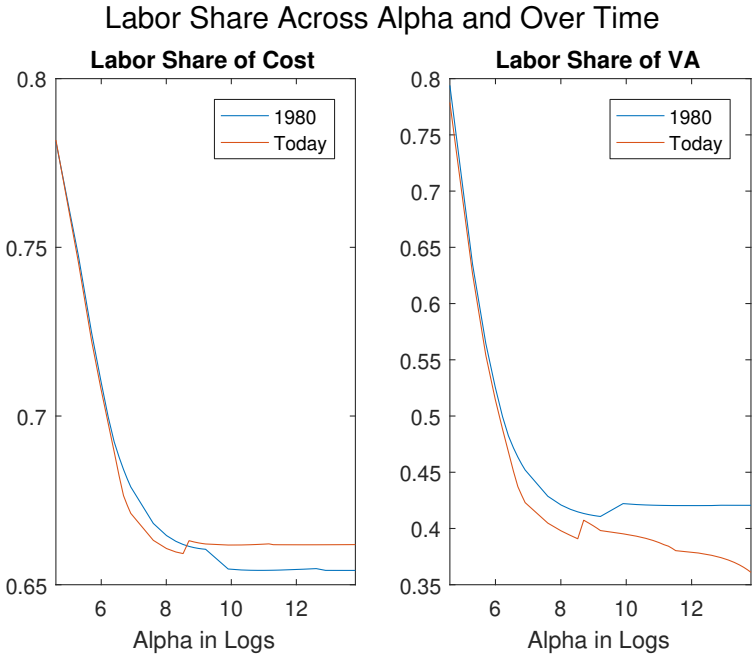


Figure 11: Simulated labor share as function of α.

Table 6 shows untargeted moments, both at the firm level and aggregate. The model delivers simulated moments that are close to the data on a variety of outcomes. Regarding the aggregate moments, revenue concentration increases and as it does the aggregate LS declines. As a consequence of the wage changes within-firms due to IT and of the employment shares the wage bill of managers increases and that of plant-workers falls. The change in wage bills comes from Eden and Gaggl (2018) who assign occupations to Routine or Non-routine labor; in the model such classification corresponds to non-managers and managers respectively. As a consequence of all the described movements the aggregate LS falls.

On the lower part of Table 6, I report firm-level moments. The capital-labor ratio in the cross-section increases with firm value added, a feature observed in the data. On the other hand, changes in wages in large firms during the last 30 years are in accordance with the results in Song et al. (2015). In particular, wages fall at the bottom and increase at the top. While the data shows larger increases at the CEO level than at the lower managerial levels, this might be due to changes in CEO compensation packages, an issue with which I do not deal.

Moment	Model	Data	Data Source
A.Aggregate moments			
Revenue Concentration, CR4 % change	45	[12, 40]	Autor et al., (2017)
Slope(log(CR4),log(Aggregate LS))	-0.26	[0.12, -0.40]	Autor et al.,(2017)
Routine Aggregate Labor Share, % change	-1.14	-42.86	Eden and Gaggl, (2017)
Non-Routine Aggregate Labor Share, % change	21.13	23.33	Eden and Gaggl, (2017)
Aggregate Labor Share	-9.14	-13.11	Karabarbounis and Neiman, (2012)
B. Firm-level moments			
Slope(log(Capital-Labor Ratio),log(VA))	0.18	*	Pending
p50 Wage, % change in top 1% firms	-6.3	-7	Bloom et al., (2017)
p75 Wage, % change in top 1% firms	10.96	64	Bloom et al., (2017)
p99 Wage, % change in top 1% firms	7.44	137	Bloom et al., (2017)

Table 6: Untargeted Moments: Data and simulated. Data on levels are from 1980 and data on changes use the 1980-2015 period; Alternatively I use the most proximate years available in the cited papers. Brackets indicate ranges of estimates in the data moments.

6.1 Counterfactual

I will conclude with an exercise to separate the role of IT capital prices and demand elasticity in the recent decline of the labor share of national income and other aggregates.

7 Conclusion

Robert E. Lucas (1978) suggests two extensions of his theory that seem "important and promising" and aimed at capturing some of the complexity observed in the hierarchies of actual business firms. He argues that "with diminishing returns on span-of-control built in at each managerial level, an equilibrium with multilevel management should consist of an allocation of capital and product workers over "level one" managers, an allocation of these units over level two managers, and so forth, until all resources are distributed over the top level managers". This is exactly what I do in this paper, where I try to highlight the "complex substitution possibilities in production". Capturing human capital considerations such as those related to managerial careers across occupations and within hierarchies is the other extension he deems important, and Caicedo et al. (2018) explores learning in such a context.

The shape of the typical workplace has changed dramatically in the last 30 years as firms have responded to the introduction of computers, email and other internal communication technologies like Enterprise Resource Planning and others. CEOs have particularly benefitted from such a trend by freeing their time from more routine decision tasks and focusing on the really exceptional decisions. In the 1950's large firms in the US occupied open office layouts with typewriters, secretaries and clerks who used their time printing, typing, and organizing paperwork in file cabinets. Today such a workplace does not exist anymore as firms use computers and software to manage. Interestingly, this manager-biased technical change has not decreased the demand for managers as it is feared will happen with unskilled workers and automation forms like robots and so on but in fact the contrary has occurred. While I abstract from other automation forces such as those in Acemoglu and Restrepo (2018), but I show that despite managerial IT capital not being directed towards the substitution of unskilled labor per se, because of it's complementarity to skilled labor it also substitutes unskilled labor and hurts unskilled wages.

If Robert Lucas was to vacation in Quebec today he would see that much of the business landscape has changed and the fewer mom-and-pop stores and restaurants have given way to franchises of large corporations much as he expected. This is not just a casual observation as in fact Autor et al. (2017) document an increasing concentration of economic activity in large firms in the last decades. In this paper I argue Information Technology capital has increased the volume and speed of problem solv-

ing that managers in business firms can deal with, increased their span of control and reduced the communication costs across-layers, particularly benefiting large firms with more complex organizations.

References

Daron Acemoglu. Why do new technologies complement skills? directed technical change and wage inequality. *Quarterly Journal of Economics*, 113, No.4:1055–1089, 1998. 5

Daron Acemoglu. Directed technical change. *Review of Economic Studies*, 113, No.4: 1055–1089, 2002. 5

Daron Acemoglu. Labor- and capital-augmenting technical change. *Journal of European Economic Association*, 1, No.1:1–37, 2003. 5

Daron Acemoglu. Equilibrium bias of technology. *Econometrica*, 75, No.5:1371–1410, 2007. 5

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 2011. 5

Daron Acemoglu and Pascual Restrepo. The race between man and machine: Implications of technology for growth, factor shares and employment. *Forthcoming, American Economic Review*, 2018. 5, 32

Daron Acemoglu, Gino Gancia, and Fabrizio Zilibotti. Offshoring, innovation and wages. *Mimeo*, 2010. 6

Anders Akerman, Ingvil Gaarder, and Magne Mogstad. The skill complementarity of broadband internet. *The Quarterly Journal of Economics*, xxxx:17811824, 2015. 7

Pol Antràs, Luis Garicano, and Esteban Rossi-Hansberg. Offshoring in a knowledge economy. *Quarterly Journal of Economics*, 121:31–77, 2006. 5

Pol Antràs, Luis Garicano, and Esteban Rossi-Hansberg. Organizing offshoring: Middle managers and communication costs. *In The Organization of Firms in a Global Economy, Elhanan Helpman, Dalia Marin, and Thierry Verdier, ed. Cambridge, MA: Harvard University Press.*, 2008. 5

D. Autor and D. Dorn. The growth of low-skill service jobs and the polarization of the u.s. labor market. *American Economic Review*, 103, No.5:1553–97, 2013. 6

David Autor. Market structure, organizational structure and wage structure. *MIT Lecture Notes 14.662*, 2015. 26

David Autor, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. The fall of the labor share and the rise of superstar firms. *ANBER Working Paper No. 23396*, 2017. 4, 21, 32

David H. Autor. The polarization of job opportunities in the u.s. labor market: Implications for employment and earnings. *Center for American Progress and The Hamilton Project*, April, 2010. 3

David H. Autor. Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, 344(6186):843851, 2014. 3

David H. Autor, Frank Levy, and Richard J. Murnane. The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333, 2003. 3, 7

Simcha Barkai. Declining labor and capital shares. *LBS Working Paper*, 2017. 6

Nicholas Bloom, Luis Garicano, Raffaella Sadun, and John Van Reenen. The distinct effects of information technology and communication technology on firm organization. *Management Science*, 60(12):2859–2885, 2014. 6

T. F. Bresnahan, E. Brynjolfsson, and L. M. Hitt. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, 117:339376, 2002. 7

Santiago Caicedo, Robert E. Lucas Jr., and Esteban Rossi-Hansberg. Learning, career paths, and the distribution of wages. *American Economic Journal: Macroeconomics (Forthcoming)*, 2018. 32

Lorenzo Caliendo and Esteban Rossi-Hansberg. The impact of trade on organization and productivity. *Quarterly Journal of Economics*, 127:3:1393–1467, 2012. 5, 25

Lorenzo Caliendo, Giordano Mion, Luca David Opromolla, and Esteban Rossi-Hansberg. Productivity and organization in portuguese firms. *CEPR DP*, 123:4: 809–852, 2015a. 5

Lorenzo Caliendo, Ferdinando Monte, and Esteban Rossi-Hansberg. The anatomy of french production hierarchies. *Journal of Political Economy*, 123:4:809–852, 2015b. 5, 8, 13, 25

Francesco Caselli and Wilbur John Coleman. The world technology frontier. *American Economic Review*, 96, No.3:499–522, 2006. 5

Jan DeLoecker and Jan Eeckhout. The rise of market power and the macroeconomic implications. *NBER WP*, Aug 2017. 4, 28

M Eden and P Gaggl. On the welfare implications of automation. *Review of Economic Dynamics*, 2018. 6, 16, 26, 30

Robert Feenstra and Gordon Hanson. The impact of outsourcing and high-technology capital on wages: estimates for the united states. *Quarterly Journal of Economics*, 114, No.3:907–940, 1999. 6

William Fuchs and Luis Garicano. Matching problems with expertise in firms and markets. *Journal of the European Economic Association*, 8, No.2-3:354–364, 2010. 5

William Fuchs, Luis Garicano, and Luis Rayo. Optimal contracting and the organization of knowledge. *Review of Economic Studies*, 82, No.2:632–658, 2015. 5

Luis Garicano. Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, 108(5):874–904, 2000. 5

Luis Garicano and Esteban Rossi-Hansberg. Organization and inequality in a knowledge economy. *The Quarterly Journal of Economics*, 121(4):1383–1435, 2006. 6

M. Goos, A. Manning, and A. Salomons. Explaining job polarization: Routine-biased technological change and offshoring. *The American Economic Review*, 104 (8):2509–2526, 2014. 6

G. M. Grossman, E. Helpman, E. Oberfield, and T. Sampson. The productivity slowdown and the declining labor share: A neoclassical exploration. *CEPR Discussion Paper 12342*, 2017. 6

Gene Grossman and Esteban Rossi-Hansberg. Trading tasks: a simple theory of offshoring. *American Economic Review*, 98, No.5:1978–1997, 2008. 6

IMF. Understanding the downward trend in labor income shares. *World Economic Outlook*, April, 2017. 3

L. Karabarbounis and B. Neiman. The global decline of the labor share. *The Quarterly Journal of Economics*, 129(1):61–103, oct 2013. 3, 4

Matthias Kehrigy and Nicolas Vincent. Growing productivity without growing wages: The micro-level anatomy of the aggregate labor share decline. *Working Paper*, 2014. 3, 4, 21, 26

Michael T. Kiley. The supply of skilled labor and skill-biased technological progress. *The Economic Journal*, 109, No.458:708–724, 1999. 5

Dongya Koh, Ral Santaeullia-Llopis, and Yu Zheng. Labor share decline and intellectual property products capital. *BGSE WP927*, 2018. 4

Danial Lashkari and Mart Mestieri. Gains from trade with heterogeneous income and price elasticities. *Harvard WP*, 2017. 26

D. Marin and T. Verdier. Globalization and the empowerment of talent. *Journal of International Economics*, 86, No.2:209–223, 2012. 5

D. Marin and T. Verdier. Corporate hierarchies and international trade: Theory and evidence. *Journal of International Economics*, 94, No.2:295–310, 2014. 5

Ezra Oberfield and Devesh Raval. Micro data and macro technology. *NBER WP20452, Econometrica R&R*, 2014. 3, 4, 22

Devesh Raval. What's wrong with capital in the 21st century's model? In Heather Boushey, J. Bradford DeLong, and Marshall Steinbaum, editors, *After Piketty: The Inequality Agenda post-Capital in the 21st Century*, chapter 4. Harvard University Press, 2017a. 4

Devesh Raval. The micro elasticity of substitution and non-neutral technology. *Revise and Resubmit, RAND Journal of Economics*, 2017b. 25, 26

Jr. Robert E. Lucas. On the size distribution of business firms. *The Bell Journal of Economics*, 9, No.2):508–523, 1978. 5, 32

Andres Rodriguez-Clare and Natalia Ramondo. Growth, size and openness: A quantitative approach. *Mimeo*, January, 2010. 6

Jae Song, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. Firming up inequality. *NBER WP21199 QJE R&R*, 2015. 21, 30

Alexandra Spitz-Oener. The impact of outsourcing and high-technology capital on wages: estimates for the united states. *Journal of Labor Economics*, 24, No.2:235–270, 2006. 6

Mathias Thoenig and Thierry Verdier. A theory of defensive skill-biased innovation and globalization. *American Economic Review*, 93, No.3:709–728, 2003. 5