

NBER WORKING PAPER SERIES

NOT A TYPICAL FIRM:
THE JOINT DYNAMICS OF FIRMS, LABOR SHARES,
AND CAPITAL–LABOR SUBSTITUTION

Joachim Hubmer
Pascual Restrepo

Working Paper 28579
<http://www.nber.org/papers/w28579>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021

We thank Daron Acemoglu, David Autor, Dirk Krueger, as well as seminar participants at the Federal Reserve Bank of Richmond, the University of Toronto, and Wharton for their comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Joachim Hubmer and Pascual Restrepo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Not a Typical Firm: The Joint Dynamics of Firms, Labor Shares, and Capital–Labor Substitution
Joachim Hubmer and Pascual Restrepo
NBER Working Paper No. 28579
March 2021
JEL No. E22,E23,E24,E25

ABSTRACT

The decline in the U.S. labor share is far from uniform across firms. While the aggregate labor share has declined, especially in manufacturing, retail, and wholesale, the labor share of a typical firm in these industries has risen. This paper studies the dynamics of the substitution of capital for labor at the firm level and its implications for market structure and labor shares. We introduce a model of firm dynamics where firms make costly upfront investments to adopt the capital-intensive technologies required to automate additional tasks. Following a decline in the price of capital, large firms automate more of their tasks and become more capital intensive; while the median firm continues to operate a more labor-intensive technology. In line with firm-level data, our model generates transitions in which the labor share of the median firm increases at the same time as the aggregate labor share declines. We use an extension of our model that allows for endogenous markups to study the role of rising competition and reallocation towards more productive and higher-markup firms as another driver of the decline in the labor share during 1982–2012. We provide a quantitative decomposition showing that reallocation played a minor role in explaining the decline in the labor share in U.S. manufacturing but an important role in retail and other sectors. The substitution of labor with cheaper capital in a widening range of tasks played a more dominant role in explaining the decline of the manufacturing labor share.

Joachim Hubmer
University of Pennsylvania
133 S 36th St
Philadelphia, PA 19104
jhubmer@sas.upenn.edu

Pascual Restrepo
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
pascual@bu.edu

One of the most striking facts of the recent past is the fall of labor’s share of GDP in the United States and many other countries.¹ After being stable throughout much of the last century, the U.S. labor share declined from a peak of 66% in the 1980s to 60% in 2014, as can be seen in Figure 1.² The decline of the labor share is not driven by changes in the industry composition of the U.S. economy. Figure 1 shows that, if anything, holding industry shares in GDP constant at their 1987 levels would have led to a more pronounced and earlier decline in the labor share.³ Indeed, most of the decline in the U.S. labor share is driven by a sizable decline dating back to the mid 80s in the share of value added accruing to labor in retail and wholesale trade, and in particular in manufacturing.

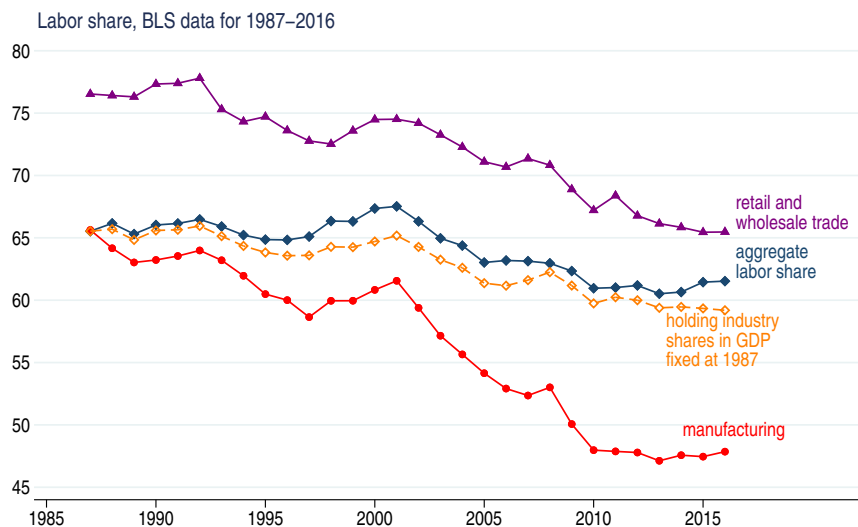


FIGURE 1: LABOR SHARE IN THE U.S. The blue line plots the headline BLS measure for the U.S. non-farm business sector. The gold line plots a counterfactual labor share holding industry shares in GDP constant at their 1987 values. The red and purple lines plot the labor shares for manufacturing and for retail and wholesale trade. Data from the BLS Multifactor Productivity Tables.

Despite growing consensus on these facts, there is an ongoing debate about the underlying causes of the decline. One set of explanations points to the increased substitution of capital for labor. In these accounts, the development of new or more efficient capital-

¹ See for instance Elsby, Hobijn and Şahin (2013); Karabarbounis and Neiman (2013); Piketty (2014); Dao, Das and Koczan (2019). For a different perspective arguing that the decline in the labor share is exclusively a U.S. phenomenon, see Gutierrez and Piton (2020).

²The figure plots the labor share in the U.S. from the BLS. Although there is some consensus about the decline in the labor share, there is more debate on the exact magnitude of the decline and how it is affected by different measurement issues such as the treatment of self-employment (Elsby, Hobijn and Şahin, 2013), housing (Rognlie, 2015), depreciation (Bridgman, 2018), income shifting by business owners (Smith et al., 2019), and intangible capital (Koh, Santaaulalia-Llopis and Zheng, 2020).

³See also Elsby, Hobijn and Şahin (2013); Hubmer (2020).

intensive technologies—as evidenced by the decline in the price of capital equipment and software—leads to the substitution of capital for labor and a decline in the labor share. This substitution can take place along an aggregate production function with an elasticity of substitution greater than one (Karabarbounis and Neiman, 2013; Eden and Gaggl, 2018; Hubmer, 2020), or within tasks in a widening range of tasks that can now be profitably automated (Acemoglu and Restrepo, 2018). The fact that the decline in the labor share is more pronounced in manufacturing, and within that sector in industries and firms adopting new automation technologies or that are more capital-intensive (Acemoglu and Restrepo, 2020; Acemoglu, Lelarge and Restrepo, 2020; Hubmer, 2020) lends some credence to these explanations.

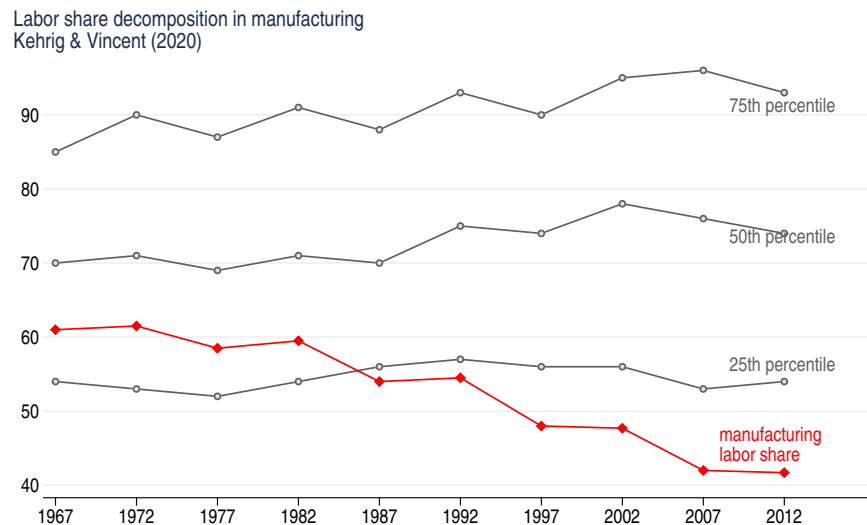


FIGURE 2: LABOR SHARE IN U.S. MANUFACTURING FIRMS. The red line plots the aggregate labor share in manufacturing. The gray lines plot the labor shares for firms at the 25th, 50th, and 75th percentile of the labor share distribution. Data from Kehrig and Vincent (2020).

This evidence notwithstanding, recent studies using firm-level data have revealed that the decline in labor shares is far from uniform across firms in an industry. While the aggregate labor share has declined, the labor share of the *typical* U.S. firm has risen (Autor et al., 2020; Kehrig and Vincent, 2020). For example, Figure 2, which we reproduced using data from Kehrig and Vincent (2020), shows that, in manufacturing, the firms located at the 25th, 50th, and 75th percentile of the labor share distribution saw a mild increase in their labor shares over time at the same time as the aggregate labor share in the sector plummeted. As shown by Autor et al. (2020), a similar pattern holds across all U.S. sectors and is also evident for other countries. At a first glance, these facts cast some doubt on explanations based on capital-labor substitution. After all, a simple

version of these theories in which firms face similar prices and have access to the same technologies would imply a more uniform decline in labor shares. Instead, these facts seem to favor a second set of explanations that emphasize the role of rising concentration and the reallocation of economic activity towards the top firms in an industry (Barkai, 2020; De Loecker, Eeckhout and Unger, 2020; Autor et al., 2020; Baqaee and Farhi, 2020b). This reallocation, which could be the result of increased competition or winner-takes-all dynamics, reduces the aggregate labor share because the top firms in each sector have higher markups. In these accounts, the decline in the labor share reflects an increase in the aggregate markup resulting from reallocation and does not represent technological changes favoring the substitution of capital for labor.⁴

This paper starts from the observation that the adoption of modern automation technologies facilitating the substitution of capital for labor is quite uneven and concentrates on large firms. Our main point is that, once we account for this form of heterogeneity in technology adoption, one can make sense of the dynamics of the labor share decline across firms and industries as a response to lower capital prices. As a result, explanations of the labor share decline based on capital–labor substitution are consistent, both qualitatively and quantitatively, with the new firm level evidence put forth by Autor et al. (2020) and Kehrig and Vincent (2020). Moreover, accounting for these differences in technology adoption leads to different conclusions regarding the relative contribution of markups and technology to the decline in the labor share across industries. In particular, we provide a series of decompositions that account for the uneven adoption of automation technologies, and show that capital-labor substitution played a key role in explaining the decline in the manufacturing labor share during 1982–2012; whereas rising competition and reallocation towards firms with lower labor shares played a key role in retail and other sectors.

We develop these points via three exercises. First, we revisit the firm labor share evidence through the lens of a standard firm-dynamics model with a CES demand structure augmented with capital–labor substitution decisions across tasks (as in Zeira, 1998; Acemoglu and Restrepo, 2018). The key assumption in our model is that firms must make costly upfront investments to adopt the capital-intensive technologies required to automate tasks. This assumption is motivated by micro evidence pointing to the skewed adoption of new capital-intensive technologies, including robotics, dedicated machinery, and specialized software. For instance, Acemoglu, Lelarge and Restrepo (2020) show that,

⁴A variant of these explanations emphasizes the role of technologies that favor larger firms (see Unger, 2019; Lashkari, Bauer and Boussard, 2019). This is different from our mechanism, which emphasizes the fact that new automation technologies have a strong factor bias and change the production function of large firms, making it less labor-intensive.

in France, only 1% of manufacturing firms purchased industrial robots between 2010 and 2015, but these firms accounted for over 20% of manufacturing value added. The *Annual Business Survey* reveals a similar fact for the U.S., where the adoption of these technologies is also highly concentrated among large firms (Zolas et al., 2020).⁵ This evidence suggest that large firms not only differ in their productivity (and potentially in their markups) but also in the extent to which their production relies on capital-intensive technologies.

Using this simple version of our model in which markups are constant, we find that, following a persistent decline in the price of capital—as the one observed in the US data since the 1980s—, the resulting dynamics of the labor share across firms match the evidence put forth by Autor et al. (2020) and Kehrig and Vincent (2020). Large and growing firms automate more of their tasks and become more capital intensive. In these firms, capital and labor become substitutes, driving the decline in the aggregate labor share. Instead, the median firm will not automate additional tasks and will continue to operate a more labor-intensive technology. For the median firm, capital and labor remain complements, explaining why the labor share rises in the typical firm.⁶ Moreover, as in the Melitz-Polanec decomposition conducted by Autor et al. (2020), the decline in the labor share is driven by a more negative covariance between the market share of a firm and their labor share, and not by the change in the unweighted average of labor shares among surviving firms. Importantly, the negative covariance term reflects the fact that automation efforts in our model are episodic and coincide with periods of rapid firm expansion in terms of value added, which aligns with the joint dynamics of changes in value added and labor shares for manufacturing firms documented by Kehrig and Vincent (2020). Our model also accounts for the positive contribution of entry and the negative contribution of exit to changes in the labor share and generates an endogenous rise in sales concentration and productivity dispersion during episodes of capital–labor substitution

⁵Dinlersoz and Wolf (2018) document a similar phenomenon for advanced manufacturing technologies using the *Survey of Manufacturing Technologies* from 1993. A series of recent papers document that the adoption of industrial robots also concentrates among the top firms in each manufacturing industry (see Koch, Manuylov and Smolka, 2019; Humlum, 2019; Bonfiglioli et al., 2020).

⁶A related mechanism going back to Houthakker (1955) emphasizes the possibility that a decline in capital prices will reallocate economic activity towards more capital-intensive firms. For example, Oberfield and Raval (2014) show that this reallocation has a small but positive contribution to the aggregate substitution between capital and labor in response to lower capital prices. More recently, Kaymak and Schott (2018) argue that this mechanism can explain up to a third of the decline in the manufacturing labor share in response to lower corporate taxes—which are isomorphic to lower capital prices. This form of reallocation to ex-ante more capital intensive firms is different from the main mechanism in our model, which emphasizes the fact that new automation technologies are unevenly adopted by large firms which then become more capital intensive.

that are quantitatively in line with empirical estimates.⁷

In the second step of our analysis, we extend our model to a non-CES demand system where demand is log-concave and markups increase with firm size—Marshall’s second law of demand.⁸ This extension allows us to quantify the effects of increased competition on concentration and the labor share. We show that one can use data on the decline in the aggregate labor share, the change in the median labor share, and the change in sales concentration in each sector of the economy to decompose the decline in the labor share into a part driven by rising competition and a part driven by capital–labor substitution. Our model-based decomposition suggests that in retail and other sectors outside of manufacturing, rising competition and the ensuing concentration of economic activity at top firms account for up to 60% of the decline in the labor share during 1982–2012. However, when looking at manufacturing this pattern reverses and capital–labor substitution explains over 90% of the decline in the labor share during this period. The reason why rising competition plays a minor role in manufacturing is that concentration in this sector increased mildly (as evident from the data in Autor et al., 2020), and that, based on the distribution of firm sales, we estimate a distribution of firm productivity that is more log-convex than but not too far from Pareto. When the productivity distribution is Pareto, the decline in the labor share due to reallocation to high markup firms and the decline in within firm markups generated by rising competition exactly offset each other.

Our quantitative exercise focuses entirely on differences in markups associated with firm size (or productivity), which is what existing explanations based on the rise of superstar firms emphasize. However, the association between size and markups is imperfect. Increased competition might reallocate economic activity to high-markup firms of any size, including smaller ones. We address this possibility in our third exercise, where we provide model-free bounds for the contribution of reallocation to high-markup firms to the decline of the labor share by sector. These bounds come at the expense of additional data requirements, and rely heavily on empirical estimates of markups and variable-input elasticities from Compustat. We first provide estimates of firm-level markups and output elasticities that account for differences in the capital-intensity of technology across

⁷While we focus on falling capital prices and automation, in our framework, globalization and increased offshoring as emphasized by Elsby, Hobijn and Şahin (2013) would have similar implications for domestic labor share and market share dynamics if offshoring decisions were likewise correlated with firm size. As we are unaware of systematic firm-level evidence on offshoring by firm size, we do not address this channel in this paper.

⁸Non-CES demand systems have been widely used in trade (see Melitz and Ottaviano, 2008; Amiti, Itskhoki and Konings, 2019; Arkolakis et al., 2018). A recent and growing literature uses non-CES demand systems to quantify the distortions introduced by markups (see Edmond, Midrigan and Xu, 2018). See also Baqaee and Farhi (2020a) for a thorough discussion on these models and Marshall’s laws.

firms. In line with the key mechanism in our model, we find that over time, the largest firms in each industry have experienced a large drop in their elasticity of output with respect to variable inputs, suggesting that their production processes are becoming less labor-intensive. On the other hand, we find positive or no changes for smaller firms. Using these empirical estimates of firm markups, we reach a conclusion that supports our quantitative model-based results: empirically, reallocation towards high-markup firms can explain at most a small share of the labor share decline in manufacturing, but plays a more important role in other sectors.⁹

The rest of the paper is organized as follows. Section 1 introduces our model for the special case of a CES demand aggregator, where markups are constant, and shows that the model matches firm-level evidence on the decline of the labor share. Section 2 considers a model with variable markups, and decomposes the decline in the labor share into a part driven by reallocation and tougher competition, and a part driven by capital-labor substitution. Section 3 provides empirical bounds on the contribution of reallocation to the decline of the labor share by sector using Compustat firm-level data. The appendix contains all proofs, as well as additional details and robustness exercises.

1 MODEL WITH CES DEMAND

1.1 Description of the model

We use a standard firm-dynamics model (as in Hopenhayn, 1992; Clementi and Palazzo, 2016), and augment it to include firms' decisions to automate tasks by adopting costly capital-intensive technologies (as in Acemoglu and Restrepo, 2018). We consider an economy in discrete time and omit time subscripts unless needed. Existing firms differ in their productivity z and in the fraction of production processes or tasks which they have automated $\alpha \in [0, 1]$. A firm of type $\theta = (z, \alpha)$ produces output $y(\theta)$ by combining a continuum of tasks indexed by x via a CES production function with task substitution

⁹ Autor et al. (2020) provide a related exercise when assessing the magnitudes of their *superstar firms*' mechanism. In particular, they use empirical estimates of markups across firms and compute the extent to which the relative expansion of high markup firms contributed to the decline of the manufacturing labor share. Their markup estimates are based on production function elasticities as in De Loecker, Eeckhout and Unger (2020), or measured as revenue over total costs. Baqaee and Farhi (2020b) conduct a similar exercise, but in their case they are interested in computing the contribution of reallocation to high-markup firms to TFP growth, which they show is equal to the contribution of reallocation to the decline in the labor share. They also provide estimates based on accounting approaches and a production function approach. We follow the production function approach, but allow production functions to vary more flexibly across firms of different sizes.

elasticity $\eta \geq 0$:

$$y(\theta) = z \cdot \left(\int_0^1 \mathcal{Y}(\theta, x)^{\frac{\eta-1}{\eta}} dx \right)^{\frac{\eta}{\eta-1}}.$$

Tasks in $[0, \alpha]$ are automated and can be produced by capital or labor; whereas non-automated tasks in $(\alpha, 1]$ must be produced by labor:

$$\mathcal{Y}(\theta, x) = \begin{cases} \psi_k(x)k(\theta, x) + \psi_\ell(x)\ell(\theta, x) & \text{if } x \in [0, \alpha] \\ \psi_\ell(x)\ell(\theta, x) & \text{if } x \in (\alpha, 1] \end{cases}$$

Here, $\mathcal{Y}(\theta, x)$ denotes the quantity of task x , and $k(\theta, x)$ and $\ell(\theta, x)$ denote the utilization of capital and labor in producing task x by a firm of type θ .

Firms face a competitive market for inputs. There is a fixed supply of labor ℓ rented to firms at a wage rate w . On the other hand, the capital used for task x is produced from the final good at a rate $q(x)$ and rented to firms at a rate $1/q(x)$.

Without loss of generality, we assume that tasks are ordered so that

$$\frac{\psi_\ell(x)}{q(x) \cdot \psi_k(x)} \text{ is increasing in } x,$$

which implies that labor has a comparative advantage at high-indexed tasks.

We start with a CES demand system. Existing firms produce differentiated varieties that are then combined via a CES aggregator to produce the final good y :

$$y = \left(\int_\theta y(\theta)^{\frac{\sigma-1}{\sigma}} m(\theta) d\theta \right)^{\frac{\sigma}{\sigma-1}},$$

where $m(\theta)$ denotes the mass of firms of type θ , and $\sigma > 1$ denotes the elasticity of substitution across varieties. In what follows, we normalize the price of y to 1.

Incumbent firms begin a period with productivity z and automation level α . They then make optimal employment and capital utilization decisions and collect profits $\pi(\theta)$. Subsequently, firms draw a fixed operating cost $c_f \cdot y$, where $c_f \sim G(c_f)$, and decide whether to continue operating.¹⁰ If they continue, they can automate additional tasks in $(\alpha, \alpha']$ at a cost $c_a \cdot y \cdot (\alpha' - \alpha)$, and draw next period's productivity level z' , which follows an

¹⁰We specify that all fixed costs are paid in units of the final good and scaled by aggregate output. This normalization ensures that the model can generate a balanced growth path if labor productivity increases uniformly.

exogenous AR1 process in logs:

$$\ln z' = \rho_z \ln z + \varepsilon,$$

where $\rho_z \in (0, 1)$, $\varepsilon \sim N(\mu_z, \sigma_z)$, and $\mu_z = -(1 - \rho_z) \cdot \sigma_z^2 / (1 - \rho_z^2)$ so that the long-run mean of z is normalized to one.¹¹

We also allow for entry of new firms. In particular, every period a unit mass of potential entrants decides whether to enter the market. Entrants draw a productivity signal z from a distribution $H(z)$ and start with a common level of automation α_0 . After observing z , α_0 , as well as the realization of the fixed operating cost c_f , entrants decide whether to pay the fixed cost and enter. After paying this cost, entrants become incumbents and face the same problem described above. Finally, we assume that automation technologies also have some exogenous diffusion through entrants, and that α_0 equals the unweighted average of α among incumbents. This is a common specification used in models of technology diffusion (for a recent example, see Perla, Tonetti and Waugh, 2021).

Both incumbents and new entrants discount the future at a constant interest rate r , which we take as exogenous. Throughout, we assume that $r > g$, where g is the growth rate of output between two consecutive periods.

The timing and treatment of entry and exit in this model follows the standard firm-dynamics literature. Our key innovation is to incorporate heterogeneity in the extent to which firms have automated their production process, and to endogenize the evolution of these decisions over time as determined by the payment of a fixed cost per task c_a . This fixed cost of automation plays a crucial role in our results and ensures that firms will automate more tasks as they grow and reach the required scale.

The assumption that automating a task entails some fixed cost is plausible and intuitive. For example, consider a car-manufacturing firm contemplating the possibility of automating welding. Besides purchasing the industrial robots required to complete these tasks ($k(\theta, x)$ in our model), the firm must also hire a team of engineers and integrators to reorganize its plant and production process, and in some cases to redesign some of their products so that they are more standardized, so that the robots can be integrated seamlessly. In the case of industrial robots, these upfront investments and payments to integrators far exceed the cost of the robot system itself (Acemoglu and Restrepo, 2020). Likewise, a firm contemplating to deploy a new software to automate its logistics and inventory management decisions must pay a fixed cost for developing the software and

¹¹In Appendix C.1, we demonstrate that our findings are not sensitive to different timing assumptions. In particular, we find similar results if we let firms observe z' before making their automation decisions.

rearranging its operations (c_a in our model) and then continue to pay a variable cost for the operation of the system ($\frac{1}{q(x)\psi_k(x)}$ in our model).¹²

Besides the costly adoption decision by firms to automate additional tasks, we also assume that these technologies diffuse gradually through the entry of new firms. This is a natural starting point and a simple way to get automation technologies to fully diffuse over time, reflecting the standardization of these production techniques and the associated organizational and productive changes required to deploy them.

1.2 Equilibrium and transitional dynamics

To save on notation, we focus on an equilibrium where firms produce all tasks in $[0, \alpha]$ with capital. This will be the relevant scenario following a reduction in the cost of capital.

The production function above implies that the unit cost for a firm of type θ is

$$(1) \quad c(\theta, w) = \frac{1}{z} \cdot (\Psi_k(\alpha) + \Psi_\ell(\alpha) \cdot w^{1-\eta})^{\frac{1}{1-\eta}}.$$

This is the usual CES price index, with the difference that the *share parameters* $\Psi_k(\alpha)$ and $\Psi_\ell(\alpha)$ are endogenous and depend on the mass of tasks that are automated:

$$\Psi_k(\alpha) = \int_0^\alpha (q(x) \cdot \psi_k(x))^{\eta-1} dx, \quad \Psi_\ell(\alpha) = \int_\alpha^1 \psi_\ell(x)^{\eta-1} dx.$$

The share of capital in cost $\varepsilon_k(\alpha, w)$ —which equals the elasticity of output with respect to capital—and the share of labor in cost $\varepsilon_\ell(\alpha, w)$ —which equals the elasticity of output with respect to labor—are then given by

$$\varepsilon_k(\alpha, w) = \frac{\Psi_k(\alpha)}{\Psi_k(\alpha) + \Psi_\ell(\alpha) w^{1-\eta}}, \quad \varepsilon_\ell(\alpha, w) = \frac{\Psi_\ell(\alpha) \cdot w^{1-\eta}}{\Psi_k(\alpha) + \Psi_\ell(\alpha) \cdot w^{1-\eta}}.$$

As these equations show, by investing in automation, firms increase the share of capital in cost and reduce the share of labor, becoming more capital-intensive.

Finally, firms charge a constant markup $\mu := \sigma/(\sigma - 1) > 1$ and earn profits

$$(2) \quad \pi(\theta, w) = \sigma^{-\sigma} (\sigma - 1)^{\sigma-1} \cdot y \cdot c(\theta, w)^{1-\sigma}.$$

We focus on perfect foresight equilibria. Given a path for investment productivities

¹²One can think of these fixed costs as intangible investment, which would contribute to firms' intangible capital (see, e.g., Corrado, Hulten and Sichel, 2009, on the increasing importance of intangibles). In our model, firms not only differ in their capital stock, but also in their ability to use capital productively for additional tasks—an intangible asset.

$q_t(x)$ and an initial distribution of firms $m_0(\theta)$, an equilibrium is given by a path for wages w_t and aggregate output y_t , as well as a path for the distribution of firms $m_t(\theta)$, such that for all $t \geq 0$:

1. the ideal-price index condition holds

$$(3) \quad \int_{\theta} \mu^{1-\sigma} \cdot c_t(\theta, w_t)^{1-\sigma} \cdot m_t(\theta) \cdot d\theta = 1,$$

2. the labor market clears

$$(4) \quad \int_{\theta} y_t \cdot \mu^{-\sigma} \cdot c_t(\theta, w_t)^{-\sigma} \cdot \frac{\partial c_t(\theta, w_t)}{\partial w_t} \cdot m_t(\theta) \cdot d\theta = \ell,$$

3. automation and exit decisions maximize the value function of incumbents

$$V_t(\theta) = \pi_t(\theta) + \int \max \left\{ 0, -c_f \cdot y_t + \max_{\alpha' \in [\alpha, 1]} \left\{ -c_a \cdot y_t \cdot (\alpha' - \alpha) + \frac{1}{1+r} \mathbb{E}[V_{t+1}(\theta') | z, \alpha'] \right\} \right\} dG(c_f),$$

where $\theta' = (z', \alpha')$ denotes the vector of technology next period,

4. entry decisions maximize the value of entrants

$$V_t^e(z) = \int \max \left\{ 0, -c_f \cdot y_t + \max_{\alpha' \in [\alpha_{0t}, 1]} \left\{ -c_a \cdot y_t \cdot (\alpha' - \alpha_{0t}) + \frac{1}{1+r} \mathbb{E}[V_{t+1}(\theta') | z, \alpha'] \right\} \right\} dG(c_f),$$

where z denotes an entrant's productivity signal, and $\alpha_{0t} \equiv \frac{\int_{\theta} \alpha \cdot m_t(\theta) d\theta}{\int_{\theta} m_t(\theta) d\theta}$,

5. starting from $m_0(\theta)$, the evolution of $m_t(\theta)$ is governed by the exogenous process for z , the endogenous process for α , and optimal entry and exit decisions of firms.

To characterize equilibrium automation decisions, it is useful to define α_t^* as the level of automation that would minimize production costs if tasks could be automated at no cost. This involves automating tasks up to the point at which the unit cost of producing a task with labor equals that of producing it with capital, which defines α_t^* as

$$(5) \quad \frac{\psi_{\ell}(\alpha_t^*)}{q_t(\alpha_t^*) \cdot \psi_k(\alpha_t^*)} = w_t.$$

The following lemma characterizes automation decisions and shows that, in line with the existence of fixed costs per task, more productive firms automate more of their tasks.

LEMMA 1 *A firm with technology (α, z) sets $\alpha' = \max\{\hat{\alpha}_t(z), \alpha\}$, where $\hat{\alpha}_t(z)$ is an increasing function of z with*

$$\lim_{z \rightarrow 0} \hat{\alpha}_t(z) = 0, \quad \lim_{z \rightarrow \infty} \hat{\alpha}_t(z) = \alpha_{t+1}^*.$$

Our first proposition characterizes the steady state of this economy. Due to the diffusion of automation technology, we have that all firms eventually achieve the same level of automation, and that this is the cost-minimizing level of automation defined above:

PROPOSITION 1 *Suppose that $q_t(x) \rightarrow q(x)$. The economy admits a stationary equilibrium where $\alpha = \alpha^*$ for all firms.*

We now turn to a characterization of the transitional dynamics in response to improvements in the efficiency with which the economy can produce capital goods $q_t(x)$. Suppose that the economy is in a steady state with $q_0(x)$ and experiences a uniform proportional increase in $q_0(x)$ by $\Delta \ln q > 0$.

To analyze the response of the economy, it is useful to define two elasticities of substitution. On the one hand, we have the elasticity of substitution between capital and labor holding the level of automation constant—the *short-run elasticity of substitution*. In our model, this coincides with the elasticity of substitution across tasks η . On the other hand, we have the elasticity of substitution between capital and labor that accounts for the endogenous shifts in technology over time captured by α^* . This *long-run elasticity of substitution* is given by

$$\eta_{LR} = \eta + \frac{\partial \ln \Psi_k(\alpha) / \Psi_\ell(\alpha)}{\partial \ln \alpha} \bigg/ \frac{\partial \ln \psi_\ell(\alpha) / (q(\alpha) \cdot \psi_k(\alpha))}{\partial \ln \alpha}.$$

Because optimal automation decisions are increasing in the wage and the level of $q(x)$ (as can be seen from (5)), we have that this long-run elasticity always exceeds η .

PROPOSITION 2 *Following a uniform proportional increase in $q_0(x)$ by $d \ln q > 0$, the economy converges to a new steady state with higher output and wages, and a higher level of automation α^* . The aggregate share of labor in costs (or in value added) changes by*

$$d \ln \varepsilon_\ell = (1 - \varepsilon_\ell) \cdot (1 - \eta_{LR}) \cdot (d \ln w + d \ln q) = \frac{1 - \varepsilon_\ell}{\varepsilon_\ell} \cdot (1 - \eta_{LR}) \cdot d \ln q.$$

At the same time, for an incumbent firm with low realizations of z along the transition,

the share of labor in costs (or in value added) changes by

$$d \ln \varepsilon_\ell(\theta) = (1 - \varepsilon_\ell(\theta)) \cdot (1 - \eta) \cdot (d \ln w + d \ln q) = \frac{1 - \varepsilon_\ell(\theta)}{\varepsilon_\ell(\theta)} \cdot (1 - \eta) \cdot d \ln q.$$

One case that will be particularly relevant for our analysis emerges when the short-run elasticity is less than 1, and at the same time, the long-run elasticity is greater than 1. In this case, Proposition 2 implies that the aggregate labor share is governed by the long-run elasticity of substitution and that it will decline as capital goods become cheaper and wages increase. At the same time, firms that do not reach a big enough scale to justify investments in automating additional tasks will keep their α fixed at its initial value. For these firms, capital and labor are complements, and their labor share rises as capital goods become cheaper and wages increase.

The previous proposition suggest that, one way to explain the dynamics of the labor share in the data, is to have a parametrization of our model where the long-run elasticity of substitution exceeds 1 and the short-run elasticity is below 1. Although this case is intuitive, and some evidence supports it as a natural starting point, our model can still generate the dynamics that we see in the data even if capital and labor are complements in the long run and $\eta_{LR} < 1$. In particular, the aggregate labor share might decline following improvements in technology that reduce the cost of capital used at marginal tasks (those around α^*), instead of uniformly across all tasks as in Proposition 2.

PROPOSITION 3 *Following an increase in $q_0(x)$ for all $x > \alpha^*$ by $d \ln q > 0$, the economy converges to a new steady state with higher output and wages, and a higher level of automation α^* . The aggregate share of labor in costs (or in value added) changes by*

$$d \ln \varepsilon_\ell = -(1 - \varepsilon_\ell) \cdot (\eta_{LR} - \eta) \cdot d \ln q + (1 - \varepsilon_\ell) \cdot (1 - \eta_{LR}) \cdot d \ln w$$

At the same time, for an incumbent firm with low realizations of z along the transition, the share of labor in costs (or in value added) changes by

$$d \ln \varepsilon_\ell(\theta) = (1 - \varepsilon_\ell(\theta)) \cdot (1 - \eta) \cdot d \ln w.$$

The proposition clarifies that an improvement in the productivity of capital at marginal tasks could reduce the aggregate labor share even if $\eta_{LR} < 1$. For example, if η_{LR} approaches 1 from below, the aggregate labor share will fall while the labor share of some typical firms will increase slightly. The reason is that these marginal changes lead to the reallocation of tasks away from labor and towards capital. At the same time, there is no

change in the price of tasks below α^* , which otherwise would reduce the share of capital (assuming $\eta < 1$). This is the key difference relative to a uniform increase in $q(x)$, which also reduces the share of infra-marginal tasks produced by capital in value added. This also clarifies that, while η_{LR} is an interesting object, it only mediates the response of the labor share to uniform changes in prices, but not to task-specific changes of the sort described in Proposition 3. Under the lens of a task model, the observed decline in the labor share in response to a decline in equipment prices can have one of two interpretations. Either this decline is uniform across tasks and $\eta_{LR} > 1$, as emphasized in much of the existing literature and in Proposition 2. Or this decline is more pronounced for new types of capital used at tasks where humans had a stronger comparative advantage, as emphasized in Proposition 3. In this last case we could have a long-run elasticity of substitution above or below 1.

1.3 Calibration and quantitative results

This section presents results from a calibrated version of our model and shows that it is able to reproduce the firm-level dynamics found in the U.S. manufacturing sector. We study the effects of a uniform decline in the price of capital goods in the main text, and leave an analysis of task-specific changes to Appendix C.3.

Initial steady state: We first discuss the calibration of the parameters governing the initial steady state of the economy. We calibrate the model under the assumption that industries were in their steady state in 1982, and that at this point in time, there was a uniform (inverse) price of capital $q(x) = q_0$, which we choose to match the observed labor share in manufacturing. This choice is motivated by the fact that the decline of the manufacturing labor share of value added starts in 1982, after being roughly constant in the decades preceding it. The period after 1982 also coincided with a pronounced decline in the price of equipment and software, which will be the main driving force behind the labor share decline in our exercise.¹³ Likewise, we set the demand elasticity to $\sigma = 11$, which generates a common markup of 1.1. Finally, we pick the dispersion of firm productivity to match the observed sales concentration in manufacturing in 1982. Specifically, we target the share of sales among the top 4 firms in each manufacturing industry reported by Autor et al. (2020), which roughly corresponds to the top 1.1% of

¹³Appendix E shows that this is also the case for most economic sectors. After having labor shares that were roughly stable for the 1947–1982 period (and in some cases until 1987), some sectors experienced a decline in their labor share of value added after this period. This Appendix also provides data for capital prices and shows a sharp decline for the 1982–2012 period (see Hubmer, 2020, for more on capital prices)

firms.¹⁴ Panel I in Table 1 summarizes the obtained parameters and the corresponding data moments that we used to calibrate them.

TABLE 1: Calibration of the CES-demand model for manufacturing

	PARAMETER		MOMENT	DATA	MODEL
<i>I. Parameters governing steady state in 1982</i>					
$\ln q_0$	Inverse capital price	-5.35	Manufacturing labor share from Kehrig and Vincent (2020)	60.1%	60.2%
σ	Demand elasticity	11.0	Aggregate markup	1.10	1.10
σ_z	Std. dev. of $\ln z$ innovations	0.094	Top 4 firms' sales share in 1982 from Autor et al. (2020)	40.0%	40.1%
<i>II. Parameters governing firm dynamics</i>					
\underline{c}_f	Minimum fixed cost	$6.0 \cdot 10^{-6}$	Entry (=exit) rate from Lee and Mukoyama (2015)	0.062	0.063
ξ	Dispersion fixed cost	0.330	Size of exiters from Lee and Mukoyama (2015)	0.490	0.491
μ_e	Entrant productivity	0.935	Size of entrants from Lee and Mukoyama (2015)	0.600	0.598
<i>III. Parameters related to the elasticity of substitution</i>					
η	Task substitution elasticity	0.4	Short-run K-L elasticity from Oberfield and Raval (2014)	0.40	0.40
γ	Comparative advantage	0.95	Long-run K-L elasticity from Hubmer (2020)	1.35	1.35

Notes: The labor share corresponds to the time average over 1967–1982 in Kehrig and Vincent (2020). The relative capital price is identified only jointly with the level of the capital productivity schedule $\psi_k(x)$, which we normalize to 1. The model equivalent to the top 4 firms' sales share refers to the top 1.1% of firms, since there are on average 364 firms per 4-digit industry in the manufacturing sector as reported in Autor et al. (2020). The annual entry rate, as well as relative sizes of entrants and exiters, are from Lee and Mukoyama (2015) and based on the Annual Survey of Manufactures. The six parameters listed in Panels I and II are all jointly calibrated to match the six corresponding moments.

The parameters governing firm dynamics in our model are taken from the literature or calibrated to match entry rates, exit rates, and firm-size distributions.¹⁵ Panel II of Table 1 lists the jointly calibrated parameters and corresponding model and data moments.

The calibration of the schedule of capital and labor productivities across tasks is more novel and important for our purposes. We adopt the following functional form:

$$\psi_\ell(x) = \left(x^{\frac{1-\eta-\gamma}{\gamma}} - 1 \right)^{\frac{1}{1-\eta-\gamma}}, \quad \psi_k(x) = 1.$$

¹⁴Although not targeted, the model roughly matches the share of sales accounted for by the top 20 firms in an industry (roughly corresponding to the top 5.5% sales share), which according to Autor et al. (2020) averages 69.7% across manufacturing industries, while the model equivalent is 68.0%.

¹⁵We follow Clementi and Palazzo (2016) in parametrizing mean and dispersion of the fixed cost distribution $G(c_f)$ to match the overall frequency and relative size of exiting firms. Specifically, we impose a Pareto distribution for the operating cost with scale parameter \underline{c}_f and tail coefficient ξ . We also match the relative size of entrants by modeling the entrant distribution $H(z)$ as a log-normal that differs from the long-run distribution of z only insofar as it has a lower mean μ_e . We fix the persistence of productivity ρ_z at 0.9, which is in the range of estimates for various TFP measures and demand shocks in Foster, Haltiwanger and Syverson (2008).

This specification implies that the short run elasticity of substitution is given by η and the long-run elasticity of substitution is constant and equal to $\eta_{LR} = \eta + \gamma$, where $\gamma > 0$ is an inverse measure of the strength of the comparative advantage of labor at higher index tasks.

We set $\eta = 0.4$ to match the estimate of the firm-level elasticity of substitution between capital and labor in Oberfield and Raval (2014), which is based on cross-regional wage variation. Although this is a reasonable starting point, we also explored the implications of using lower values for η in Appendix C.2. The reason is that wage differences across regions might persist, which implies that their estimates also account for some changes in the allocation of tasks and should be seen as an upper bound on the elasticity of substitution between tasks η . On the other hand, we set $\gamma = 0.95$ to match the available estimates of the aggregate long-run elasticity of substitution between capital and labor in response to a fall in capital prices, which place it around 1.35 (see Karabarbounis and Neiman, 2013; Hubmer, 2020).¹⁶ As emphasized in Proposition 3 and demonstrated in Appendix C.3, we can also generate the observed behavior of the manufacturing labor share through a decline in the price of capital at higher-indexed tasks even if the long-run elasticity of substitution η_{LR} is less than 1.

Transitional dynamics: We now explore the adjustment of the economy following a decrease in the price of capital. In particular, we ask if our model can generate (i) the observed decline in the manufacturing labor share at the same time as (ii) an increase in the median labor share in response to a uniform increase in $q(x)$. To this end, we calibrate the decline in the price of capital and the fixed cost of automation by targeting (i) and (ii). Intuitively, the decline in the price of capital controls the aggregate labor share, and the fixed cost of automation controls the behavior of the median labor share.

Column 1 in Table 2 summarizes the manufacturing data, and Column 2 displays the inferred shocks as well as model moments corresponding to our benchmark calibration. Column 2 shows that our model with CES demand and a fixed cost of automation per task of 0.23 generates the observed trends in (i)–(ii) as the response to a decline in the price of capital of 140 log points.¹⁷ Thus, the observation that the labor share has risen for the

¹⁶Whether the strategy pursued by Oberfield and Raval (2014) identifies a *long-run* capital–labor elasticity or a *short-run* task-level elasticity of substitution depends on the extent to which the regional wage differences used for identification are permanent. Presumably, some of the variation in regional wages is more transitory than the time-series variation in capital prices exploited in Karabarbounis and Neiman (2013); Hubmer (2020), which could reconcile these divergent estimates in the literature.

¹⁷We assume that the capital price declines at a constant rate over the sample period 1982–2012; subsequently, it remains constant.

TABLE 2: Transitional dynamics and decomposition of the manufacturing labor share using the CES demand system (1982–2012)

	DATA	MODEL	
		BENCHMARK	NO FIXED COST OF AUTOMATION
	(1)	(2)	(3)
<i>I. Parameters and inferred aggregate shocks</i>			
$d \ln q$.	1.40	1.40
c_a	.	0.23	0
<i>II. Targeted moments, 1982–2012</i>			
Δ aggregate labor share	−0.178	−0.176	−0.198
Δ median labor share	0.030	0.026	−0.198
<i>III. Other moments, 1982–2012</i>			
Δ Top 4 firms’ sales share	0.060	0.043	−0.015
Δ Top 20 firms’ sales share	0.052	0.070	−0.022
Δ log productivity dispersion	0.050	0.057	0.000
<i>IV. Melitz–Polanec decomposition from Autor et al. (2020)</i>			
Δ aggregate labor share	−0.185	−0.176	−0.198
Δ unweighted mean	−0.002	0.032	−0.198
Exit	−0.055	−0.003	0
Entry	0.059	0.006	0
Covariance term	−0.187	−0.212	0
<i>V. Covariance decomposition from Kehrig and Vincent (2020)</i>			
Market share dynamics	0.047	0	0
Labor share by size dynamics	−0.043	−0.036	0
Joint dynamics	−0.232	−0.143	0

Notes: Column (2) reports the findings from our benchmark model, which calibrates a uniform decline in the capital price (log-linear over 1982–2012) as well as the automation fixed cost to replicate both the change in the aggregate as well as in the median labor share (as documented by Kehrig and Vincent (2020) for the manufacturing sector). Column (3) displays a counterfactual economy with costless automation. The change in industry concentration in Panel III is from Autor et al. (2020, Table 1) and refers to the average change 1982–2012 across 4-digit manufacturing industries. The model equivalent is the top 1.1%, respectively top 5.5%, firm sales share. The change in the standard deviation of log productivity (log unit cost) is from Decker et al. (2020, Figure 3a), and computed analogously as the difference between the 2000s and 1980s. Panel IV reproduces the Melitz–Polanec decomposition from Autor et al. (2020, Table 4 Panel B), reported as the sum of consecutive 5-year changes 1982–2012. Panel V reproduces the covariance decomposition from Kehrig and Vincent (2020, Figure 5), conducted for a balanced sample of firms and one long change 1982–2012.

typical manufacturing firm is not inconsistent with capital–labor substitution playing a predominant role in the decline in the manufacturing labor share. Through the lens of our model, this observation simply tells us that the median firm operates less capital-intensive technologies than larger firms.

The fixed cost of automation required to rationalize the data is small and of a reasonable magnitude. The calibrated value of c_a implies that increasing α by an amount that decreases a firm’s labor share by one percentage point incurs a fixed cost equal to 0.6–1.2%

of annual average firm sales. In fact, in response to lower capital prices in our model, total spending on automation fixed costs as a share of manufacturing output rises from 0 in the initial steady state to a peak of 0.7% in 2005. The fixed cost of automating tasks can be thought of as an investment in R&D required to design and integrate automation equipment or software. Appendix E compares this spending to the observed behavior of total R&D expenditures in the data. Because not all R&D spending is linked to automation, we see this series as providing an upper bound on expenditures required to develop and integrate the software and equipment required for automation. In line with this view, we find that the model-implied spending on automation fixed costs amounts to only 25% of the R&D share in the data, and that the rise in spending on fixed costs aligns with the increasing share of R&D observed over time since the 80s. This comparison shows that one can make sense of the manufacturing data with a small and reasonable value for the fixed cost of automation, and that with this cost, our model does not generate a counterfactual level of R&D expenditures.

Why doesn't the typical firm automate its production processes further even though the cost of doing so is small? This is because, despite its sizable effect on factor shares, the productivity gains from automation might be small. In particular, the productivity gains from automating a task can be approximated by the product of the cost-saving gains generated by switching to produce a task with capital, given by

$$\text{cost saving gains} = \ln\left(\frac{w}{\psi_\ell(\alpha)}\right) - \ln\left(\frac{1}{q(\alpha) \cdot \psi_k(\alpha)}\right) \geq 0,$$

multiplied by the share of the automated tasks in value added—an extension of Hulten's theorem to our environment.¹⁸ This logic implies that, as α approaches the cost-minimizing level α^* , the productivity gains from automating these tasks will be positive but small. Moreover, the formula for the cost-saving gains shows that, even for smaller values of α , the productivity gains from automating tasks might be small when the comparative advantage of labor in higher-index tasks is not too strong—i.e., when γ is large as in our calibration.

On the other hand, the decline in the price of capital of 140 log points over the period 1982–2012 is in the ball park of what we observe for computer-powered equipment and

¹⁸ In particular, Appendix A shows that, as $\alpha \rightarrow \alpha^*$ from below, the TFP gains of increasing automation by $d\alpha > 0$ are given by

$$\text{TFP gains} = \left[\ln\left(\frac{w}{\psi_\ell(\alpha)}\right) - \ln\left(\frac{1}{q(\alpha) \cdot \psi_k(\alpha)}\right) \right] \cdot \frac{(w/\psi_\ell(\alpha))^{1-\eta}}{\Psi_k(\alpha) + \Psi_\ell(\alpha) \cdot w^{1-\eta}} \cdot d\alpha \geq 0.$$

software, whose price declined by 170 log points during this period (see Eden and Gaggl, 2018; Hubmer, 2020, and Appendix E for more on capital prices). Because industries also use other types of capital, we should think of this 170 log point decrease as an upper bound on the decline in the price of capital that one can feed into the model.

Figure 3 displays the aggregate and the median labor share in the model. The aggregate labor share declines steeply and comes close to the eventual (lower) steady state value by 2012. On the other hand, the median—summarizing the behavior of a typical firm—increases during the first part of the transition, and only declines below its initial level after 2012. Eventually, automation technologies diffuse through entry and as firms cycle through high- z states, which slowly drives the labor share of the typical firm down, taking several years to converge to the new steady state.

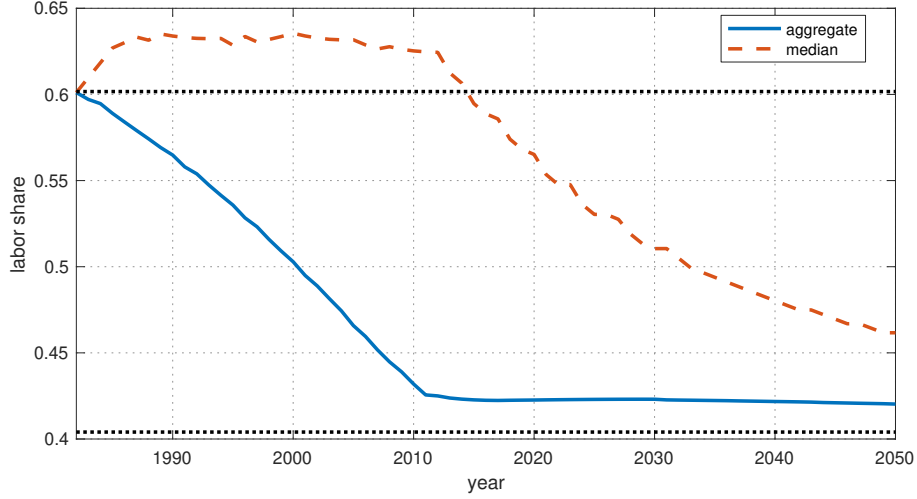


FIGURE 3: TRANSITIONAL DYNAMICS FOR THE AGGREGATE AND TYPICAL LABOR SHARES IN THE CES MODEL FOR MANUFACTURING. The horizontal dotted lines refer to the initial (upper) and final (lower) steady states of the aggregate labor share. The solid blue line plots the behavior of the aggregate labor share. The dashed red line plots the behavior of the median of firms' labor shares.

To further investigate the predictions of our model for firms' labor share dynamics and contrast them with the data, we decompose the decline in the manufacturing labor share using a Melitz–Polanec decomposition:

$$\begin{aligned}
 \Delta s_\ell & && \text{(Change in manufacturing labor share)} \\
 = \Delta \bar{s}_\ell & && \text{(Change in unweighted means)} \\
 + \omega_X(s_{\ell S} - s_{\ell X}) + \omega'_E(s'_{\ell E} - s'_{\ell S}) & && \text{(Contribution of exit and entry)} \\
 + \Delta \sum_f (\omega_f - \bar{\omega}) \cdot (s_{\ell f} - \bar{s}_\ell) & && \text{(Change in covariance).}
 \end{aligned}$$

Here, Δs_ℓ denotes the change in the manufacturing labor share over a given period of

time. This can be decomposed into the change in the unweighted mean of labor shares among continuing firms, $\Delta \bar{s}_\ell$; two terms accounting for the contributions of firms that exit the market and firms that enter the market; and the change in the covariance among continuing firms, f , between their share of value added, ω_f , and their labor share, $s_{\ell f}$. The contribution of firms that exit the market is given by their share of value added in the baseline period, ω_X , multiplied by the difference in the average labor share of continuing firms, $s_{\ell S}$, and firms that exit, $s_{\ell X}$. The contribution of firms that enter the market is given by their share of value added in the end period, ω'_E , multiplied by the difference in the average labor share of firms that enter, $s'_{\ell E}$, and continuing firms, $s'_{\ell S}$.

We follow Autor et al. (2020) and conduct this decomposition using 5-year differences, and report the sum for each component over the first 30 years of the transition in our model, corresponding to 1982–2012 in the data. Panel IV in Table 2 reproduces the decomposition from Autor et al. (2020) for manufacturing and reports the decomposition from our model. As in the manufacturing data, in our model economy the covariance term fully accounts for the aggregate decline in the labor share. Exit and entry exhibit the same qualitative patterns as in the data: on average, both entering and exiting firms have labor shares that are higher than those of surviving firms. However, the differences between these and continuing firms is not as pronounced as in the data, where many of the firms that enter or exit have labor shares that exceed 1, presumably reflecting other elements of the life cycle of firms that are not included in our model.¹⁹ The results in Panel IV show that our theory of capital–labor substitution at the task level and with fixed costs per task is capable of reproducing the new firm-level facts put forth by Autor et al. (2020). They also imply that the Melitz–Polanec decomposition is not sufficient to discriminate between explanations for the decline in the labor share based on technology or others based on rising competition and reallocation.

The dominant role of the covariance term warrants further inspection. In an account-

¹⁹In our model, a firm that enters in t starts with $\alpha = \alpha_{0t}$ and is allowed to automate prior to production. In contrast, the labor share of entering firms would be higher (amplifying the contribution of entry) if automation technology diffused more slowly and new entrants were not allowed to automate in the period they enter. Likewise, the contribution of entry and exit to the labor share would be amplified if fixed costs were more labor-intensive than the final good (our model assumes they are equally intensive). In this case, both firms that enter and exit would have higher labor shares, reflecting the importance of the labor-intensive fixed costs for small firms.

ing sense, changes in the covariance term can be decomposed as

$$\begin{aligned}
& \Delta \sum_f (\omega_f - \bar{\omega}) \cdot (s_{\ell f} - \bar{s}_\ell) && \text{(change in covariance)} \\
& = \sum_f \Delta(\omega_f - \bar{\omega}) \cdot (s_{\ell f} - \bar{s}_\ell) && \text{(market share dynamics)} \\
& + \sum_f (\omega_f - \bar{\omega}) \cdot \Delta(s_{\ell f} - \bar{s}_\ell) && \text{(labor share by size dynamics)} \\
& + \sum_f \Delta(\omega_f - \bar{\omega}) \cdot \Delta(s_{\ell f} - \bar{s}_\ell) && \text{(cross dynamics).}
\end{aligned}$$

That is, we could have a decrease in the covariance driven by a reallocation of value added towards firms with lower labor shares at baseline (the “market share dynamics” term); a more pronounced reduction in the labor share of large firms (the “labor share by size dynamics” term); or the possibility that firms that reduce their labor shares expand at the same time (the “cross dynamics” term). In practice, decomposing changes in the covariance into these terms might be challenging, since measurement error in value added will inflate the contribution of the cross dynamics term. Though we do not yet have a definite answer about the importance of these different components, Kehrig and Vincent (2020) provide some evidence suggesting that, in manufacturing, the cross dynamics have played a crucial role in explaining the decline in the labor share. In particular, using a balanced sample of firms for 1982–2012, Kehrig and Vincent (2020) find that the cross dynamics contributed -23.2 percentage points to the decline in the covariance term, accounting for 102% of the total decline within the balanced sample, while labor share dynamics by size account for 19% of the decline.

Panel V in Table 2 provides the contributions of the three components above implied by the decomposition in Kehrig and Vincent (2020). The panel also shows that our model economy broadly aligns with their data.²⁰ In particular, in our model, the covariance term is negative because large firms have a more pronounced decline in their labor share (labor share dynamics by size account for 20% of the total decline), but primarily because growing and expanding firms are the ones investing in automation and reducing their labor share at the same time (cross dynamics account for the remaining 80%). This follows directly from the policy functions provided in Lemma 1, which show that firms go through episodes of automation when they receive a series of positive productivity shocks causing their desired automation level $\hat{\alpha}_t(z)$ to exceed their current automation level α .

²⁰We compute the model moments in this panel exactly as in Kehrig and Vincent (2020), reporting the cumulative change over 30 years in a balanced sample. Therefore, both in model and data, the three terms in Panel V do not exactly add up to the covariance term in Panel IV.

For comparison, Column 3 in Table 2 summarizes the transitional dynamics of our model that would result if firms did not face any fixed cost of automating additional tasks. We find that while the aggregate labor share evolves similarly in this counterfactual economy, the firm-level dynamics of labor shares and market shares are strongly at odds with the data. The comparison between Columns 2 and 3 underscores the importance of fixed costs and shows that even a small fixed cost can substantially alter the type of dynamics that we see in response to falling capital prices.

Productivity dispersion and sales concentration: Our model generates endogenous changes in productivity dispersion and sales concentration along the transition of the economy to its new steady state. For the 1982–2012 period, uneven automation choices contribute to an increase in productivity dispersion of close to 10 log points and 4 log points for the 1982–2000 period. This is broadly in line with the evidence in Decker et al. (2020), who estimate an increase in TFP dispersion by about 5 log points for the U.S. manufacturing sector, comparing the 1980s to the 2000s. In turn, widening productivity differentials generate an endogenous increase in sales concentration. In terms of concentration ratios, Autor et al. (2020) document increases of 6 percentage points for the share of sales by the top 4 firms in each manufacturing industry and 5.2 percentage points for the share of sales by the top 20 firms. The model generates comparable increases of 4.3 percentage points for the top 4 firms and 7.0 percentage points for the top 20 firms. In sum, our model endogenously generates the observed increase in sales concentration and productivity dispersion in manufacturing, and by doing so provides an alternative explanation for the correlation between lower labor shares and higher sales concentration observed across U.S. industries (see for example Barkai, 2020; Autor et al., 2020). However, our model also predicts that these differences will revert over time as automation technologies slowly diffuse to all firms, as shown in Figure 4.

2 MODEL WITH SIZE-DEPENDENT MARKUPS

This section extends our model to allow for differences in markups across firms. This extension allows us to decompose the decline in the labor share and the rise in sales concentration into a component driven by increasing competition and another driven by capital–labor substitution.

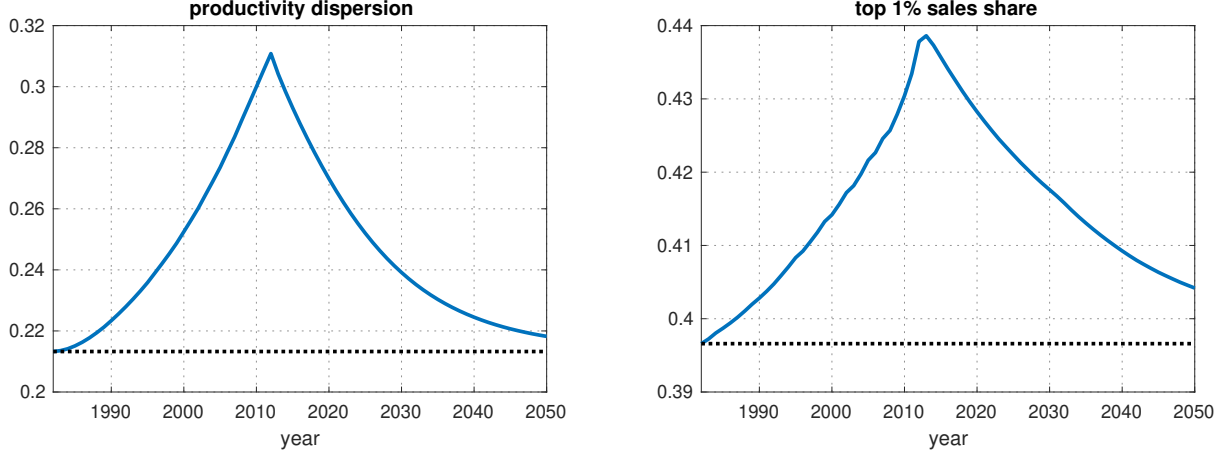


FIGURE 4: TRANSITIONAL DYNAMICS FOR PRODUCTIVITY DISPERSION AND SALES CONCENTRATION IN THE CES MODEL FOR MANUFACTURING. Productivity dispersion is computed as the cross-sectional standard deviation of log unit costs. Sales concentration is measured by the top 1% sales share. The dotted line depicts the steady state values for both variables (the distributions of relative firm productivity and relative firm sales are unchanged across steady states).

2.1 Description of the model

To allow for differences in markups, we assume that the demand faced by firms is given by a Kimball aggregator (see Kimball, 1995), where total output in an industry is defined implicitly as the solution to

$$\int_{\theta} \lambda \cdot H\left(\frac{y(\theta)}{\lambda \cdot y}\right) m(\theta) d\theta = 1.$$

Here, H is an increasing and concave function. Normalizing the price of the final good to 1 yields the demand curve faced by firms as

$$y(\theta) = y \cdot \lambda \cdot D\left(\frac{p(\theta)}{\rho}\right),$$

where the demand function D is decreasing and given by the inverse function of $H'(x)$, ρ is an endogenous summary measure of competitors' prices, which we will define below, and $p(\theta)$ is the price charged by a firm of type θ . Finally, λ is an exogenous proxy for market size, which will serve to model increasing competition.

When $H(x) = x^{1-1/\sigma}$, we obtain the typical CES demand system. The demand function $D(p)$ is simply given by the log-linear function $p^{-\sigma}$, and the competitors' price index ρ coincides with the price of the final good, which we normalized to 1. However, for general aggregators H , the demand function might take different forms and the competitors' price index—determining how other prices affect a firm's demand—will be different from the

price of the final good.

Given a path for investment productivities $q_t(x)$, a path for market size λ_t , and an initial distribution of firms $m_0(\theta)$, an equilibrium is given by a path for wages w_t , output y_t , the competitors' price index ρ_t , the distribution of firms $m_t(\theta)$ and their pricing decisions, such that:

1. pricing decisions by firms $p_t(\theta)$ maximize

$$\pi_t(\theta) = \max_p y_t \cdot \lambda_t \cdot D\left(\frac{p}{\rho_t}\right) \cdot (p - c_t(\theta, w_t)),$$

2. the ideal-price index condition holds, which defines the final good as the numeraire,

$$\int_{\theta} \lambda_t \cdot p_t(\theta) \cdot D\left(\frac{p_t(\theta)}{\rho_t}\right) m_t(\theta) d\theta = 1,$$

3. the implicit definition of competitors' price index solves

$$\int_{\theta} \lambda_t \cdot H\left(D\left(\frac{p_t(\theta)}{\rho_t}\right)\right) m_t(\theta) d\theta = 1,$$

4. the labor market clears

$$\int_{\theta} y_t \cdot \lambda_t \cdot D\left(\frac{p_t(\theta)}{\rho_t}\right) \cdot \frac{\partial c_t(\theta, w_t)}{\partial w_t} \cdot m_t(\theta) \cdot d\theta = \ell,$$

5. automation and exit decisions maximize the value function of incumbents

$$V_t(\theta) = \pi_t(\theta) + \int \max \left\{ 0, -c_f \cdot y_t + \max_{\alpha' \in [\alpha, 1]} \left\{ -c_a \cdot y_t \cdot (\alpha' - \alpha) + \frac{1}{1+r} \mathbb{E}[V_{t+1}(\theta') | z, \alpha'] \right\} \right\} dG(c_f),$$

where $\theta' = \{z', \alpha'\}$ denotes the vector of technology next period,

6. entry decisions maximize the value of entrants

$$V_t^e(z) = \int \max \left\{ 0, -c_f \cdot y_t + \max_{\alpha' \in [\alpha_{0t}, 1]} \left\{ -c_a \cdot y_t \cdot (\alpha' - \alpha_{0t}) + \frac{1}{1+r} \mathbb{E}[V_{t+1}(\theta') | z, \alpha'] \right\} \right\} dG(c_f),$$

where z is the initial productivity of an entrant,

7. starting from $m_0(\theta)$, the evolution of $m_t(\theta)$ is governed by the exogenous process for z , the endogenous process for α , and optimal entry and exit decisions of firms.

2.2 Equilibrium and Rising Competition

As in our baseline model, the equilibrium of the economy converges to a steady state where $\alpha = \alpha^*$. Moreover, firms' automation decisions follow similar policy functions.

We first characterize the behavior of markups—the new element in the extended model. Denote the price charged by a firm with unit cost c as a function of costs by $p_t^*(c)$, the resulting markup by $\mu_t^*(c)$, the quantity of goods sold by $y_t^*(c)$, and total sales by $s_t^*(c)$.

We will impose two assumptions known as Marshall's weak and strong second law:²¹

$$\begin{aligned} \text{(Marshall's weak second law)} \quad & -\frac{D'(x)}{D(x)}x \text{ is greater than 1 and increasing in } x \\ \text{(Marshall's strong second law)} \quad & x + \frac{D(x)}{D'(x)} \text{ is positive and log-concave} \end{aligned}$$

PROPOSITION 4 *Under Marshall's weak second law, firms with lower costs c charge lower prices $p_t^*(c)$ but higher markups $\mu_t^*(c)$. Moreover, under Marshall's strong second law, markups and prices, $\mu_t^*(c)$ and $p_t^*(c)$, are a log-convex function of costs, which implies lower passthroughs for more productive firms. Finally, sales $s_t^*(c)$ are a log-concave and decreasing function of costs.*

The weak second law requires that, as firms lower their prices, their demand becomes more inelastic. This implies that larger and more productive firms will charge lower prices but set a higher markup. The strong second law requires marginal revenue to be positive and log-concave, which ensures that more productive firms have lower passthroughs. Figure 5 illustrates these features by plotting the behavior of markups, prices, and sales. For comparison, the figure also illustrates the behavior of these equilibrium objects in a CES demand system, where markups are constant and prices and sales are log-linear functions of cost.

We impose Marshall's second weak and strong law for three reasons. First, these assumptions receive support from the data (see for example Baqaee and Farhi, 2020a). Second, this demand structure offers a tractable way of capturing the type of pricing dynamics that one would get in oligopolistic competition models. In those models, larger firms recognize that an increase in their price will have a disproportionate effect on the industry price index, and so they recognize that their revenue is less responsive to changes in their price. Third, and more importantly, these assumptions ensure that more productive firms are larger but charge higher markups, and that an increase in competition will

²¹Throughout, we say that a function $y = f(x)$ is log-concave if $\ln y$ is concave in $\ln x$.

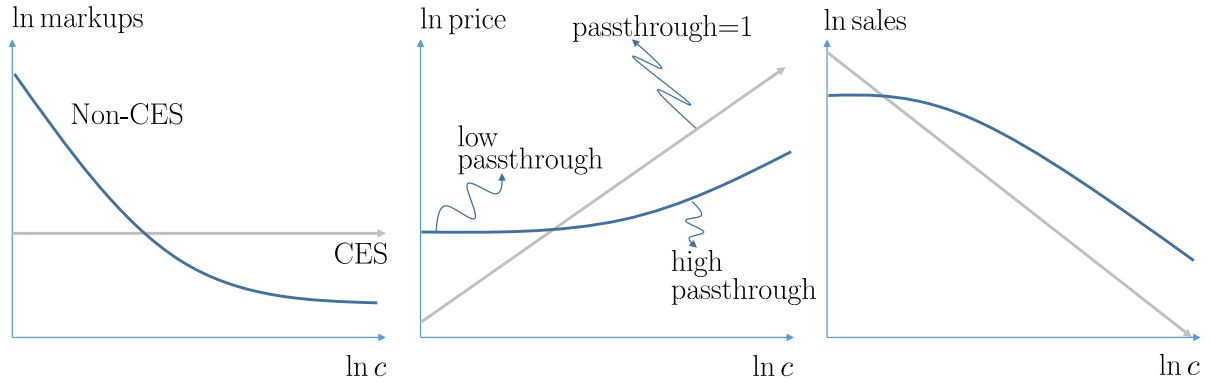


FIGURE 5: MARKUPS, PRICES AND SALES. The figure presents the log of markups, prices and sales as a function of firm unit costs in a CES demand system and a non-CES system where demand satisfies Marshall's weak and strong second laws.

result in a reallocation of economic activity towards these large firms with high markups. This is precisely the reasoning behind theories emphasizing how rising competition can reduce the labor share.

To capture the effects of rising competition, we consider an increase in market size λ . The next proposition shows that increases in λ lead to tougher competition and a reallocation towards more productive firms. We will focus on a stationary equilibrium, and with some abuse of notation, denote by $x(z)$ the equilibrium value for object x associated with a firm with productivity z . Finally, we denote the share of sales of a firm by $\omega(z)$. Recall that in a stationary equilibrium all firms have the same automation level and we don't have to keep track of this dimension.

PROPOSITION 5 *An increase in λ at time t has the following effects in the stationary equilibrium time:*

- all firms will reduce their markups, $\mu(z)$;
- for any two firms with $z > z'$, we have that $\mu(z)/\mu(z')$ is decreasing;
- for any two firms with $z > z'$, we have that $\omega(z)/\omega(z')$ is increasing.

The proposition shows that, as market size increases, firms will be pushed towards the more elastic segments of their demand curves and will respond by reducing their markups. However, the reduction in markups is not uniform: larger and more productive firms will reduce their markups and relative prices disproportionately, capturing a greater share of the market-size increase. This is a direct consequence of Marshall's strong second

law, which implies that large and more productive firms have smaller passthroughs. In particular, as tougher competition for workers increases real wages, large firms respond via a modest increase in their prices (and a large reduction in their markup) while small firms respond via a more sizable increase in their prices (and a smaller reduction in their markups).

This response by firms generates an ambiguous contribution of markups to the labor share of the economy. On the one hand, firms reduce their markups, which contributes to an increase in the labor share. On the other hand, there is a reallocation of economic activity from small firms with low markups to large firms with high markups, which contributes to a decline in the labor share.

PROPOSITION 6 *Suppose that the distribution of productivities in a stationary equilibrium is given by $z \sim f(z)$. The aggregate labor share is given by*

$$s_\ell = \varepsilon_\ell(\alpha^*, w^*) \cdot \frac{1}{\mu},$$

where the aggregate markup μ is a sales weighted harmonic mean of firm-level markups:

$$\frac{1}{\mu} := \int_z \frac{1}{\mu(z)} \cdot \omega(z) \cdot f(z) \cdot dz$$

An increase in λ increases the aggregate markup if the distribution of productivity z is log-convex (i.e., more convex than Pareto), lowers it if the distribution of productivity z is log-concave (i.e., less convex than Pareto), and leaves it unchanged if the distribution of productivity z is log-linear (i.e., Pareto).²²

The proposition shows that the effects of rising competition on the aggregate markup depend on the distribution of productivity across firms. This insight is well known and recognized in the literature (see for instance Melitz and Ottaviano, 2008; Autor et al., 2020). As we will show in our quantitative exercises, this insight is also relevant for understanding the calibrated effects of rising competition on the labor share and the aggregate markup.

Propositions 5 and 6 are stated for a stationary equilibrium, where all firms have the same automation level α . When firms differ in α —for example, along the transition following a decline in capital prices—, rising competition will also reallocate economic activity towards more capital-intensive firms (those with larger values of α), generating an additional force towards reducing the labor share. The reason is that these firms

²²We refer to a distribution as log convex (log concave) if its PDF $f(z)$ is log convex (log concave).

experience less cost inflation as competition tightens and firms bid up wages. Although we do not have sharp theoretical results for this general case, we will show that this interaction between differences in technology and rising competition plays an important role in some of our quantitative exercises.

In the next section, we will quantify these effects and also explain how they interact with automation decisions in determining the behavior of the labor share.

2.3 Calibration

The previous section characterized the properties of the equilibrium of our extended model and some of the effects of an increase in market size and competition. We now explore the full transitional dynamics in response to rising competition and lower capital costs. This allows us to analyze how capital–labor substitution and reallocation interact in determining the dynamics of markups, the labor share of firms, and the aggregate labor share.

We start by calibrating the model to the U.S. manufacturing sector. We first explain how we parametrize and calibrate the structural parameters. We then explain how we use the calibrated model to infer the decline in the price of capital $1/q$ and the increase in competition λ required to match the observed behavior of the labor share and sales concentration in this sector.

Manufacturing: We first discuss the calibration of the parameters governing the initial steady state of the economy. We calibrate the model under the assumption that industries were in their steady state in 1982, and that at this point in time, there was a uniform (inverse) price of capital $q(x) = q_0$, which we choose to match the observed aggregate labor share in manufacturing.

Following Edmond, Midrigan and Xu (2018), we parametrize H using the specification from Klenow and Willis (2016), which satisfies Marshall’s weak and strong second laws. The key property from this specification is that the demand elasticity faced by a firm setting a price p is given by

$$(6) \quad \text{demand elasticity}(p) = \sigma \cdot D(p/\rho)^{-\frac{\nu}{\sigma}},$$

which decreases as p falls—so that more productive firms face more inelastic demand. In this specification, σ controls the average demand elasticity faced by firms, and the *supra-elasticity* ν/σ controls the extent to which markups rise for more productive firms as well

as the extent to which these firms have lower passthroughs. If $\nu = 0$, the demand system simplifies to the standard CES aggregator. The full specification for H and the derivation of equation (6) are provided in Appendix B. We will pay special attention to calibrating ν/σ , which will be the key parameter determining the net effect of rising competition on sales concentration.

In our general model, we can no longer assume that productivity z is log-normal as we did in the CES version. The reason is that, as shown in Proposition 4, sales are a log-concave function of costs (and hence productivity), and so we wouldn't be able to match the observed levels of sales concentration with a log-normal distribution of productivities. In fact, because the sales distribution is approximately Pareto, we need to entertain the possibility that the productivity distribution is more convex than Pareto. To allow for this possibility, we assume that productivity is determined by a latent factor \tilde{z} that follows an AR(1) process as before. This latent factor then determines productivity as:

$$z = \exp\left(F_{Weibull(n,\zeta)}^{-1}(\Phi(\tilde{z}))\right),$$

where Φ denotes the Gaussian cdf, and $F_{Weibull(n,\zeta)}^{-1}$ the inverse CDF of a Weibull random variable with shape parameter $n > 0$ and scale parameter $\zeta > 0$. This specification implies that the log of z will follow a Weibull distribution whose CDF is given by

$$F_{Weibull(n,\zeta)}(x) = 1 - e^{-\left(\frac{x}{\zeta}\right)^n}.$$

The Weibull distribution generalizes the exponential distribution by introducing the shape parameter n , which controls the log-convexity of the distribution. In the benchmark case with $n = 1$, the density of $\ln z$ is log-linear, or equivalently, the limit distribution for z is Pareto with tail index ζ . For $n < 1$, the density of $\ln z$ is log-convex, or equivalently, the limit distribution for z is more convex than Pareto. As shown in Proposition 6, the implications of an increase in competition on the aggregate markup depend crucially on whether the z -distribution is log-convex, log-concave, or log-linear. Thus, we will pay special attention to calibrating n , which will be another key parameter determining the net effect of rising competition on the aggregate markup and the labor share.

We jointly calibrate the parameters governing markups $\{\sigma, \nu\}$, and the parameters governing the behavior of productivity $\{\zeta, n\}$ to match the aggregate markup, the difference between the aggregate and the median labor share, and the share of sales among the top 4 and top 20 firms in manufacturing in 1982. Panel I in Table 3 summarizes the moments used and the resulting parameters. Although these parameters are jointly

calibrated with the remaining model parameters, one can see how the calibration of each of them is informed by these different moments. We calibrate $\sigma = 6.1$ to match an aggregate markup of 1.15, which is in line with our empirical estimates in Section 3. More importantly, we estimate a supra-elasticity ν/σ of 0.22 to match the difference between the labor share of the median firm in 1982 and the aggregate labor share. Because we assume that the economy is in a steady state initially, the difference in labor shares is entirely due to markups, which in turn reflect changes in the curvature of demand between the median and more productive firms.²³

TABLE 3: Steady state calibration of the non-CES demand model: Manufacturing

	PARAMETER		MOMENT	DATA	MODEL
<i>I. Parameters governing steady state in 1982</i>					
$\ln q_0$	Inverse capital price	-6.55	Aggregate labor share	60.1%	60.2%
σ	Demand elasticity	6.10	Aggregate markup	1.15	1.15
ν/σ	Demand supra-elasticity	0.22	Median labor share ratio	1.169	1.101
ζ	Weibull scale	0.086	Top 20 firms' sales share	69.7%	69.7%
n	Weibull shape	0.78	Top 4 firms' sales share	40.0%	40.0%
<i>II. Parameters governing firm dynamics</i>					
\underline{c}_f	Minimum fixed cost	$4.6 \cdot 10^{-6}$	Entry (=exit) rate	0.062	0.063
ξ	Dispersion fixed cost	0.310	Size of exiters	0.490	0.488
μ_e	Entrant productivity	0.876	Size of entrants	0.600	0.601

Notes: The aggregate and median labor share correspond to the respective time averages in the manufacturing sector 1967–1982 in Kehrig and Vincent (2020); the median is displayed as ratio over the aggregate. The two concentration measures are from Autor et al. (2020) and correspond to the manufacturing sector in 1982. The model equivalents refer to the top 1.1% and top 5.5% of firms ranked by sales (since there are on average 364 firms per 4-digit manufacturing industry). The data moments and calibrated parameters in Panel II follow the model with CES demand, see Table 1. The eight parameters in Panel I and II are jointly calibrated to match the eight corresponding moments.

For the parameters governing the productivity distribution, we calibrate $\{\zeta, n\}$ to match the share of sales by the top 4 firms and the top 20 firms in each manufacturing industry, which roughly corresponds to the top 1.1% and top 5.5% of sales. Intuitively, a higher top 5.5% share of sales indicates a thicker tail of the productivity distribution—a

²³ The Klenow–Willis specification used in our analysis limits the amount of markup variation that one can generate for a given dispersion in sales. As a result, when matching the dispersion in sales, our model generates two thirds of the observed ratio of the median vs. the aggregate labor share in manufacturing (1.169) in steady state. Picking the higher estimate for average markups of 1.15 partly mitigates this problem and allows us to be conservative, in the sense that it helps us generate enough dispersion in markups by firm size. Indeed, our calibrated demand supra-elasticity (0.22) is squarely in between the benchmark (0.14) and high supra-elasticity (0.30) calibration of Edmond, Midrigan and Xu (2018), who choose their benchmark estimate to match labor share dispersion by firm size in U.S. data. Thus, our model generates variation in markups by firm size that aligns with the data, which is reassuring since this is the key dimension of heterogeneity governing the effects of reallocation on the labor share in response to tougher competition in this model. Instead, the targeted difference between the median and the mean might reflect other sources of markup and labor share variation unrelated to firm size.

higher ζ . Moreover, conditional on the top 5.5% share, a higher top 1.1% share signals lower values of n , which indicate a more than proportional increase in productivity quantiles as we move to the top of the sales distribution. By targeting these two moments we find that a value of $n = 0.78$ —a small deviation from Pareto and thus a moderate degree of log-convexity—fits the sales data. The fact that we estimate a log-convex distribution for productivity follows from our theory. As the above introduced demand system implies that sales are less dispersed than productivity, this is needed to match the roughly log-linear distribution of sales observed in the data.

All of the remaining parameters are calibrated in the same way as in the CES model, although we obtain different estimates due to the different demand structure. Panel II in Table 3 summarizes the targeted moments and resulting parameter estimates.

Retail: We pursue an identical calibration approach for retail and summarize the resulting parameters in Table 4. The main difference relative to what we found for manufacturing is that for retail we calibrate a lower value of $n = 0.54$ (implying more log-convexity of the z -distribution), which is necessary to match the high sales share of the top 4 firms in each retail industry (the top 0.023%) of 15% in 1982 relative to the (also high) share of the top 20 firms (the top 0.12%) in each retail industry of 30%.

TABLE 4: Steady state calibration of the non-CES demand model: Retail

	PARAMETER		MOMENT	DATA	MODEL
<i>I. Parameters governing steady state in 1982</i>					
$\ln q_0$	Inverse capital price	-7.35	Aggregate labor share	70.4%	70.5%
σ	Demand elasticity	9.0	Aggregate markup	1.15	1.15
ν/σ	Demand supra-elasticity	0.20	Median labor share ratio	1.169	1.106
ζ	Weibull scale	0.023	Top 20 firms' sales share	29.9%	29.9%
n	Weibull shape	0.54	Top 4 firms' sales share	15.1%	15.1%
<i>II. Parameters governing firm dynamics</i>					
\underline{c}_f	Minimum fixed cost	$5.2 \cdot 10^{-7}$	Entry (=exit) rate	0.062	0.062
ξ	Dispersion fixed cost	0.250	Size of exiters	0.490	0.488
μ_e	Entrant productivity	0.868	Size of entrants	0.600	0.599

Notes: The aggregate labor share corresponds to the BLS Multifactor Productivity Tables estimate for the retail sector. The ratio median-to-aggregate is from Kehrig and Vincent (2020), which refers to manufacturing, since in the retail sector the census data used, e.g., by Autor et al. (2020) does not allow to compute the labor share of value added. The two concentration measures are from Autor et al. (2020) and correspond to the retail sector in 1982. The model equivalents refer to the top 0.023% and top 0.116% of firms ranked by sales (since there are on average 17,259 firms per 4-digit retail industry). The data moments and calibrated parameters in Panel II follow the model with CES demand, see Table 1. The eight parameters in Panel I and II are jointly calibrated to match the eight corresponding moments.

2.4 Quantifying the role of competition and capital–labor substitution

We now use the calibrated version of the model to decompose the observed decline in the labor share and the rise of sales concentration in manufacturing into a decline in the price of capital and an increase in competition. That is, we use the model to obtain the underlying change in q and λ required to match the observed decrease in the aggregate labor share and the observed rise in concentration in this sector from 1982 until 2012. Using the shocks recovered by the model, we can then provide a decomposition of the observed changes into a part explained by each shock and their interaction. We will also provide results for the retail sector that show that our model is capable of capturing a wide range of configurations for the relative importance of these shocks, depending on the concentration of sales in a sector, the rise in sales concentration over time, and the observed decline in the labor share.

Our exercise puts the capital–labor substitution and rising competition explanations on equal footing and uses the model to infer the contribution of these two driving forces. As we will see, our model provides different results across sectors, suggesting that our approach is capable of discriminating between these two forces.²⁴

The key assumption behind this approach is that there are no other forces affecting sales concentration or reducing the labor share of an industry. In particular, our exercise ignores the potential role of labor market power and other forms of rising market power that could contribute to rising markups and rising concentration via a channel different from an increase in market size λ .²⁵

Manufacturing: Given data on changes in (i) the aggregate labor share, (ii) the median labor share, and (iii) concentration, we can use our model to infer (a) the q -shock, (b) the λ -shock, and (c) the fixed cost of automation. Column 1 in Table 5 summarizes the manufacturing data and Column 2 provides our results. We find that, to match the observed trends in (i)–(iii), our model requires a large decline in the price of capital of 140 log points, a mild increase in competition of 4%, and a fixed cost of automating tasks of 0.16—that is, about 0.5% of average annual revenue in manufacturing industries for a one percentage point reduction in the labor share. The reason why we calibrate a small

²⁴In Appendix E, we discuss the interpretation of the inferred changes in q and λ further and relate them to empirical proxies.

²⁵Berger, Herkenhoff and Mongey (2019) find that labor market concentration has decreased over time. Using a quantitative model, they estimate that this reduction in labor market power by itself would have raised the labor share by 3 percentage points over the time period considered in our analysis.

increase in competition is that the rise in sales concentration in the sector has been mild. Indeed, manufacturing is the sector with the lowest increase in the share of sales accruing to the top 20 firms from 1982–2012 according to the data in Table 1 of Autor et al. (2020). Given the small increase in competition, the model requires a large decline in the price of capital to match the observed behavior of the labor share. As before, a small fixed cost of automating tasks is enough to ensure a rise in the labor share of the median firm.

We find that, in response to these two shocks, our model provides a good fit to the manufacturing data. Panels IV and V of Table 5 describe the dynamics of the labor share across firms by employing the Melitz–Polanec and the covariance decomposition introduced in Section 1.3. In line with the data, we still find a crucial role for a decline in the covariance between firm sales and their labor share in explaining the aggregate decline in the labor share, with most of this change driven by firms that expand at the same time as their labor share declines (the cross-dynamics term).

To understand the contribution of each of these shocks, we provide a series of counterfactual scenarios where we shut them down sequentially. In Column 3, we shut down the increase in competition. We find that capital–labor substitution explains 16.3 of the observed 17.8 percentage points decline in the manufacturing labor share. Moreover, capital–labor substitution explains 3.8 percentage points of the observed 6.0 percentage points increase in sales concentration among the top 4 firms in the sector. In line with these findings, the results in Column 4 show that, when we shut down the decrease in the price of capital, the increase in competition does not contribute materially to the decline in the manufacturing labor share. Its main role is to increase sales concentration among the top 4 firms by 1.9 percentage points.

Several factors combine to explain why rising competition plays a small role in manufacturing. First, given the modest increase in sales concentration, most of which can be accounted for by the fact that larger firms will be the ones involved in capital–labor substitution, we calibrate a small increase in competition of $d \ln \lambda = 4\%$. Second, the distribution of productivities in manufacturing is close to Pareto, which implies that rising competition does not have a substantial effect on the labor share. Finally, the large decline in the manufacturing labor share requires a sizable decline in the price of capital, which then becomes the dominant force in this sector. As we will see in the case of retail, different observed trends can lead to different conclusions about the relative importance of these shocks.

As a final exercise, Column 5 shuts down both shocks, which serves to illustrate the type of labor share dynamics across firms that the model generates in a steady state. By

TABLE 5: Transitional dynamics and decomposition of the manufacturing labor share using a non-CES demand system (1982–2012)

	MODEL				
	DATA (1)	BENCHMARK (2)	ONLY EFFECTS OF $d \ln q$ (3)	ONLY EFFECTS OF $d \ln \lambda$ (4)	NO SHOCKS (5)
<i>I. Parameters and inferred aggregate shocks</i>					
$d \ln q$.	1.40	1.40	0	0
$d \ln \lambda$.	0.04	0	0.04	0
c_a	.	0.16	0.16	0.16	0.16
<i>II. Targeted moments, 1982–2012</i>					
Δ aggregate labor share	−0.178	−0.173	−0.163	−0.002	.
Δ median labor share	0.030	0.026	0.014	0.002	.
Δ Top 4 firms’ sales share	0.060	0.060	0.038	0.019	.
<i>III. Other moments, 1982–2012</i>					
Δ Top 20 firms’ sales share	0.052	0.091	0.066	0.022	.
Δ log productivity dispersion	0.050	0.073	0.061	0.000	.
<i>IV. Melitz–Polanec decomposition from Autor et al. (2020)</i>					
Δ aggregate labor share	−0.185	−0.173	−0.163	−0.002	.
Δ unweighted mean	−0.002	0.041	0.033	0.003	.
Exit	−0.055	−0.012	−0.013	−0.015	.
Entry	0.059	0.010	0.012	0.010	.
Covariance term	−0.187	−0.212	−0.195	0.000	.
<i>V. Covariance decomposition from Kehrig and Vincent (2020)</i>					
Market share dynamics	0.047	0.062	0.063	0.063	0.071
Labor share by size dynamics	−0.043	0.018	0.016	0.066	0.071
Joint dynamics	−0.232	−0.254	−0.240	−0.130	−0.135
<i>VI. Markups, 1982–2012</i>					
Δ log aggregate markup	.	0.011	0.010	0.001	.
Within firm change in markup	.	−0.024	−0.020	−0.011	.
Reallocation	.	0.035	0.030	0.011	.

Notes: Column (2) reports the findings from our benchmark model, which calibrates a uniform decline in the capital price (over 1982–2012), an increase in competition (over 1982–2012), as well as the automation fixed cost to replicate the change in the aggregate as well as in the median labor share (as documented by Kehrig and Vincent (2020)), and the increase in the top 4 firms' sales share (Autor et al., 2020, Table 1), all for the manufacturing sector. Column (3) shows results when shutting down the competition shock, column (4) when shutting down instead the price of capital shock, and column (5) shows results when shutting down the competition (stationary equilibrium). The change in the standard deviation of log productivity (log unit cost) is from Decker et al. (2020, Figure 3a), and computed analogously as the difference between the 2000s and 1980s. Panel IV reproduces the Melitz–Polanec decomposition from Autor et al. (2020, Table 4 Panel B), reported as the sum of consecutive 5-year changes 1982–2012. Panel V reproduces the covariance decomposition from Kehrig and Vincent (2020, Figure 5), conducted for a balanced sample of firms and one long change 1982–2012. Panel VI displays the log change in the aggregate markup, as well as a decomposition into within firm and reallocation components as in Baqaee and Farhi (2020b).

construction, there are no changes in the aggregate labor share, sales concentration, and other aggregates. Instead, we focus on Panel V, which summarizes the results regarding the covariance decomposition, and illustrates the behavior of firm dynamics in terms of sales and value added along a steady state. Even in a steady state, all the components of the change in the covariance term are different from zero. In particular, the model generates a positive contribution of the market size and the labor share by size dynamics of equal size. The reason is that there is mean reversion in productivity: firms that are large today and charge high markups will be small tomorrow, explaining the positive contribution of market size dynamics. Likewise, firms that are large today have low labor shares and are expected to contract and lower their markups as they do so, explaining the positive contribution of the labor share by size dynamics. These two effects are mostly offset by the reduction in the labor share of firms that receive positive productivity shocks and expand, and as a result reduce their markups, explaining the negative contribution of the cross term.²⁶

The findings in Column 5 suggest that covariance decompositions must be interpreted with caution before drawing inferences about the forces behind the decline in the labor share. As we have seen, even in the absence of shocks we will have non-zero components that cancel each other out. By comparing the covariance decomposition in Column 5 with Column 1, we see that the main features of the data that require an explanation are the fact that the labor share by size dynamics are negative and not positive, and that the cross dynamic term is more negative than what one would observe in a steady state. Columns 2 and 3 shows that, at a qualitative level, these are precisely the sort of dynamics that we get in response to a decline in the price of capital.

Panel VI of Table 5 summarizes the predictions of our model for markups, which cannot be directly observed in the data. The labor share in an industry can be written as

$$s_\ell = \frac{\varepsilon_\ell}{\mu},$$

where ε_ℓ is the share of labor in costs for the whole industry and μ is the aggregate

²⁶These three numbers do not add up to zero because these moments are computed for incumbents. Likewise, there are some ergodic movements for other objects in the other panels of the table, which in all cases reflect differences between incumbents and firms that enter and exit along a steady state. However, these movements are an order of magnitude smaller than what we see in the data and not relevant for our purposes, so we do not report them nor discuss them here.

industry markup, defined as the harmonic-sales-weighted mean of markups across firms:

$$\frac{1}{\mu} = \sum_f \omega_f \cdot \frac{1}{\mu_f}.$$

As suggested by this decomposition and also by Proposition 6, this is the relevant notion of an aggregate markup.²⁷ In particular, the above decomposition shows that an industry labor share might decrease because of technology or changes in factor prices—captured by the share of labor in costs—or because of an increase in the industry markup.

Our model predicts a mild increase in the manufacturing markup of $d \ln \mu = 1.1\%$ (from 1.15 to 1.162). However, as anticipated in Proposition 6, this net effect masks two distinct forces. On the one hand, firms lower their markups in response to rising competition. The contribution of these within-firm changes is given by

$$\text{within-firm changes} := \sum_f \omega_f \cdot \Delta \ln \mu_f,$$

which lowered the manufacturing markup by 2.4% during this period (in this expression, the sum is over all continuing firms). On the other hand, rising competition generates a reallocation of output towards firms with higher markups, which improves allocative efficiency. The contribution of reallocation to markups is then given by

$$\text{allocative changes} := \Delta \ln \mu - \sum_f \omega_f \cdot \Delta \ln \mu_f,$$

which increased the manufacturing markup by 3.5% during this period. As emphasized by Baqaee and Farhi (2020b), allocative changes also measure the contribution of improvements in allocative efficiency to sectoral TFP.²⁸ Thus, our model predicts that rising competition and capital–labor substitution among large firms improved the allocative efficiency of the manufacturing sector, and that as a whole, these improvements raised manufacturing TFP by 3.5%.

²⁷ This decomposition follows from the chain of identities

$$s_\ell = \frac{\sum_f s_{\ell f} y_f}{\sum_f y_f} = \frac{\sum_f s_{\ell f} y_f}{\sum_f \frac{1}{\mu_f} y_f} \frac{\sum_f \frac{1}{\mu_f} y_f}{\sum_f y_f} = \varepsilon_\ell \cdot \frac{1}{\mu}.$$

In the last step, we used the fact that $\sum_f s_{\ell f} y_f$ equals the wage bill and $\sum_f \frac{1}{\mu_f} y_f$ equals total cost in the whole industry.

²⁸ An application of Proposition 1 in their paper to our context shows that the contribution of changes in allocative efficiency to TFP are given by $-\sum_f \omega_f \Delta \ln \mu_f - \varepsilon_\ell \Delta \ln s_\ell - \varepsilon_k \Delta \ln s_k$, where s_ℓ and s_k denote the share of labor and capital in industry value added and ε_ℓ and ε_k denote cost shares. With constant returns to scale, and after some algebra, one can show that $-\varepsilon_\ell \Delta \ln s_\ell - \varepsilon_k \Delta \ln s_k = d \ln \mu$.

Retail: We now conduct the same model-based decomposition for the U.S. retail sector. Table 6 summarizes the retail data and presents the inferred shocks and their implications.²⁹ In retail, we observe a decline in the labor share of 10.2 percentage points and a rise in sales concentration among the top 4 firms of 14 percentage points. This represents a vast increase in sales concentration, given that the top 4 firms represent only 0.023% of all firms in a retail industry, on average. For comparison, in manufacturing, the top 4 firms represent 1.1% of the firms in each industry, and their share of sales only increased by 6 percentage points. Our model can explain the observed patterns for retail with a decline in the price of capital of 75 log points and an increase in competition of 41%—an order of magnitude larger than in manufacturing. As before, a small fixed cost of automating tasks is enough to ensure a rise in the labor share of the typical firm.

The counterfactual scenarios in Columns 3 and 4 suggest an important role for both the relative price of capital as well as the competition shock in explaining the observed trends in retail. Moreover, the results point to a significant interaction between these two shocks. For example, the estimated q -shock by itself causes the aggregate labor share to decline by 4.2 pp, while the estimated λ -shock generates a 2.0 pp decline. Yet, in combination the two shocks generate a decline of 10.3 pp. Thus, their interaction accounts for 4.1 percentage points—40% of the total decline. Two mechanisms are responsible for this interaction. First, rising competition reallocates activity towards more automated firms, since these are the firms who experience less cost inflation as competitors bid up wages. This form of reallocation also contributes to the decline in the labor share. Second, rising competition implies that more productive firms will now account for a greater share of sales, generating extra incentives for automation among these firms and giving a higher weight to those firms in determining the aggregate labor share.

Turning to the predicted behavior of markups, our model predicts an increase in the aggregate markup in retail of 3.3% (from 1.15 to 1.19). This is the result of a 2.5% decrease in the within-firm component and a 5.8% increase driven by improvements in allocative efficiency. Thus, our model predicts that rising competition improved allocative efficiency in retail, and this generated an increase in TFP in that sector of 5.8%.

Other sectors: We also conducted a similar decomposition for other economic sectors, including wholesale, as well as utilities & transportation sectors.³⁰ We relegate the de-

²⁹For retail, our best proxy for the behavior of the labor share of the typical firm is the change in the mean unweighted payroll share (rather than the median labor share) reported by Autor et al. (2020). For aggregates in this sector, we target aggregate labor share data from the BLS, both in levels and changes.

³⁰We omit the finance and services sectors. In the former, measuring the labor share of valued added is conceptually difficult, while the latter did not experience a decline in its labor share.

TABLE 6: Transitional dynamics and decomposition of the retail labor share using a non-CES demand system (1982–2012)

	MODEL			
	DATA (1)	BENCHMARK (2)	ONLY EFFECTS OF $d \ln q$ (3)	ONLY EFFECTS OF $d \ln \lambda$ (4)
<i>I. Parameters and inferred aggregate shocks</i>				
$d \ln q$.	0.75	0.75	0
$d \ln \lambda$.	0.41	0	0.41
c_a	.	0.15	0.15	0.15
<i>II. Targeted moments, 1982–2012</i>				
Δ aggregate labor share	−0.102	−0.103	−0.042	−0.020
Δ unweighted mean labor share	0.044	0.042	−0.023	0.014
Δ Top 4 firms' sales share	0.140	0.115	0.001	0.108
Δ Top 20 firms' sales share	0.163	0.204	0.003	0.188
<i>III. Other moments, 1982–2012</i>				
Δ log productivity dispersion	.	0.014	0.001	0.000
<i>IV. Markups, 1982–2012</i>				
Δ log aggregate markup	.	0.033	0.001	0.029
Within firm change in markups	.	−0.025	−0.011	−0.024
Reallocation	.	0.058	0.012	0.053

Notes: Column (2) reports the findings from our benchmark model, which calibrates a uniform decline in the capital price (over 1982–2012), an increase in competition (over 1982–2012), as well as the automation fixed cost to replicate the change in the aggregate labor share in retail (BLS), the unweighted mean labor share change among surviving firms (Autor et al., 2020, Table 5), and the increase in the top 4 as well as top 20 firms' sales share (Autor et al., 2020, Table 1), all for the retail sector. Column (3) shows results when shutting down the competition shock, and column (4) when shutting down instead the price of capital shock. Panel IV displays the log change in the aggregate markup, as well as a decomposition into within firm and reallocation components as in Baqaee and Farhi (2020b).

tails behind the calibration and the data used to Appendix D. Figure 6 summarizes our findings and for reference contrasts them with our results for manufacturing and retail. In wholesale as well as utilities & transportation, the estimated q -shock by itself accounts for roughly two thirds of the aggregate labor share decline, while the remaining one third is almost entirely due to the interaction of the two shocks, which reinforce each other.

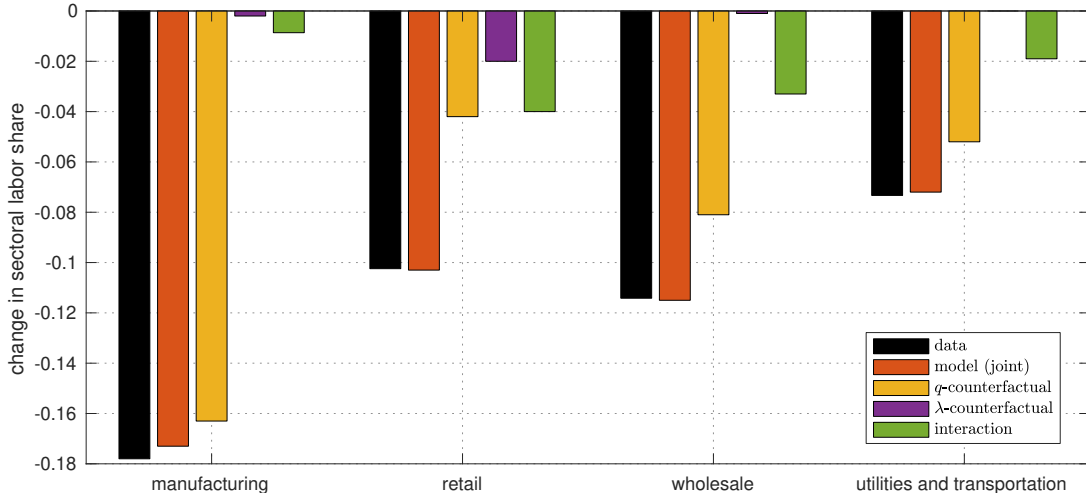


FIGURE 6: MODEL-BASED DECOMPOSITION OF SECTORAL LABOR SHARE CHANGES. For each sector, this figure shows the observed change in the aggregate labor share, the benchmark model with q - and λ -shocks jointly calibrated to match observed labor share and concentration changes, the two counterfactuals restricting to a single shock at a time, and the residual interaction term. The data sources are Kehrig and Vincent (2020) for manufacturing, and BLS MFP Tables for all other sectors. See also Tables 5, 6 and 10 for more detailed information .

We conclude that while the fall in the manufacturing labor share is almost entirely attributed to automation through the lens of our model-based decomposition exercise, in non-manufacturing sectors there is an important role for rising competition and the way in which it interacts with differences in capital-labor substitution across firms.

3 BOUNDING THE EFFECTS OF REALLOCATION AND CHANGES IN MARKUPS ON THE LABOR SHARE

Our model-based decomposition accounts for the role of rising competition generating a reallocation towards the most productive firms, which also happen to be the ones with largest markups. However, the relationship between productivity (or size) and markups might not be perfect, and one could also imagine that rising competition reallocates economic activity towards firms with high markups that are not necessarily large or the most productive. This form of reallocation would also contribute to a rise in the aggregate

markup and a decline in the labor share.

In this section, we explore the role of this more general form of reallocation empirically. The greater generality comes at the expense of requiring estimates of markups at the firm level. In what follows, we explain our approach to estimating markups and compare them with the existing literature. We then use our estimates to conduct several accounting exercises designed to illustrate the contribution of more general forms of reallocation and changes in markups to the labor share of manufacturing and other economic sectors.

3.1 Output elasticities by firm size

We estimate markups for a firm f in industry i at time t as

$$\mu_{ft} = \frac{\varepsilon_{vft}}{s_{vft}},$$

where ε_{vft} is the elasticity of gross output quantities with respect to the quantity of variable inputs used, and s_{vft} denotes the share of variable inputs in the firm's revenue.

The elasticity ε_{vft} is not observed and must be estimated from firm-level data on revenue (y), expenditures in variable inputs (v), capital (k), and investment (x). Following Olley and Pakes (1996) and Akerberg, Caves and Frazer (2015), we make the following assumptions:

- A1 differences across firms in the price of variable inputs reflect quality, which implies that we can treat expenditures in variable inputs as a measure of their quality-adjusted quantity
- A2 revenue is given by a revenue production function of the form

$$\ln y_{ft} = z_{ft} + \varepsilon_{vc(f)t}^R \cdot \ln v_{ft} + \varepsilon_{kc(f)t}^R \cdot \ln k_{ft} + \epsilon_{ft},$$

where $c(f)$ denotes groups of firms with a common technological bias and the same process for their revenue productivity, and ϵ_{ft} is an i.i.d. ex-post shock orthogonal to k_{ft} and v_{ft}

- A3 unobserved productivity z_{ft} evolves according to a Markov process of the form

$$z_{ft} = g(z_{ft-1}) + \zeta_{ft},$$

where ζ_{ft} is orthogonal to k_{ft} and v_{ft-1} , and

A4 the gross output production function exhibits constant returns to scale in capital and variable inputs, which implies that quantity elasticities are given by

$$(7) \quad \varepsilon_{v,ft} = \varepsilon_{vc(f)t}^R / \left(\varepsilon_{vc(f)t}^R + \varepsilon_{kc(f)t}^R \right).$$

Assumptions A1–A3 are standard in the literature. Assumption A4 is added to deal with the fact that we don’t observe prices, and so the estimation procedure yields revenue elasticities, not the quantity elasticities that are relevant for computing markups (Bond et al., 2020).³¹

Under these assumptions, and given a grouping of firms $c(f)$, we can estimate revenue elasticities following the usual approach from Akerberg, Caves and Frazer (2015), which uses investment as a proxy variable for unobserved productivity. This requires a first-stage regression where we first compute “true” output as

$$\ln \tilde{y}_{ft} = \mathbb{E}[\ln y_{ft} | \ln x_{ft}, \ln k_{ft}, \ln v_{ft}, t, c(f)] = h(\ln x_{ft}, \ln k_{ft}, \ln v_{ft}; \theta_{c(f)t}^h).$$

Here $\theta_{c(f)t}^h$ is a parametrization for a flexible function h that might vary over time and between groups of firms. For any pair of revenue elasticities $\varepsilon_{vc(f)t}^R$ and $\varepsilon_{kc(f)t}^R$, one can then compute

$$\tilde{z}_{ft} = \ln \tilde{y}_{ft} - \varepsilon_{vc(f)t}^R \cdot \ln v_{ft} - \varepsilon_{kc(f)t}^R \cdot \ln k_{ft},$$

estimate the flexible model

$$\tilde{z}_{ft} = g(\tilde{z}_{ft-1}; \theta_{c(f)t}^g) + \tilde{\zeta}_{ft},$$

where $\theta_{c(f)t}^g$ is a parametrization for a flexible function g , and form the following moment conditions that identify the revenue elasticities:

$$\mathbb{E}[\zeta_{ft} \otimes (\ln k_{ft}, \ln v_{ft-1})] = 0.$$

³¹ Suppose that revenue is given by $y = p(q)q$, where $p(q)$ is the inverse demand curve. Quantity elasticities and revenue elasticities are then linked according to

$$\varepsilon_v^R = \left(\frac{p'(q)q}{p(q)} + 1 \right) \cdot \varepsilon_v \quad \varepsilon_k^R = \left(\frac{p'(q)q}{p(q)} + 1 \right) \cdot \varepsilon_k,$$

where $1/\mu = \left(\frac{p'(q)q}{p(q)} + 1 \right)$. Moreover, firms will demand variable inputs until $\varepsilon_v^R = s_v$. Assuming constant returns to scale implies that $\varepsilon_v = \varepsilon_v^R / (\varepsilon_v^R + \varepsilon_k^R)$, as wanted.

This approach requires the choices of variable inputs to be correlated over time, which we view as a reasonable requirement.

We implement this approach using data from Compustat. Appendix F describes our sample selection and definitions. We parametrize the functions h and g using quadratic polynomials and conduct our estimation over 10-year rolling windows. More importantly, and in line with the emphasis in our model that large and growing firms might operate different technologies and have different output elasticities, we group firms by quintiles of sales in each industry. Thus, our estimation provides output elasticities that vary over time, by industry, and by quintiles of firm size in each industry. This represents a significant deviation from previous papers which assume that all firms in a given industry share a common bias of technology.

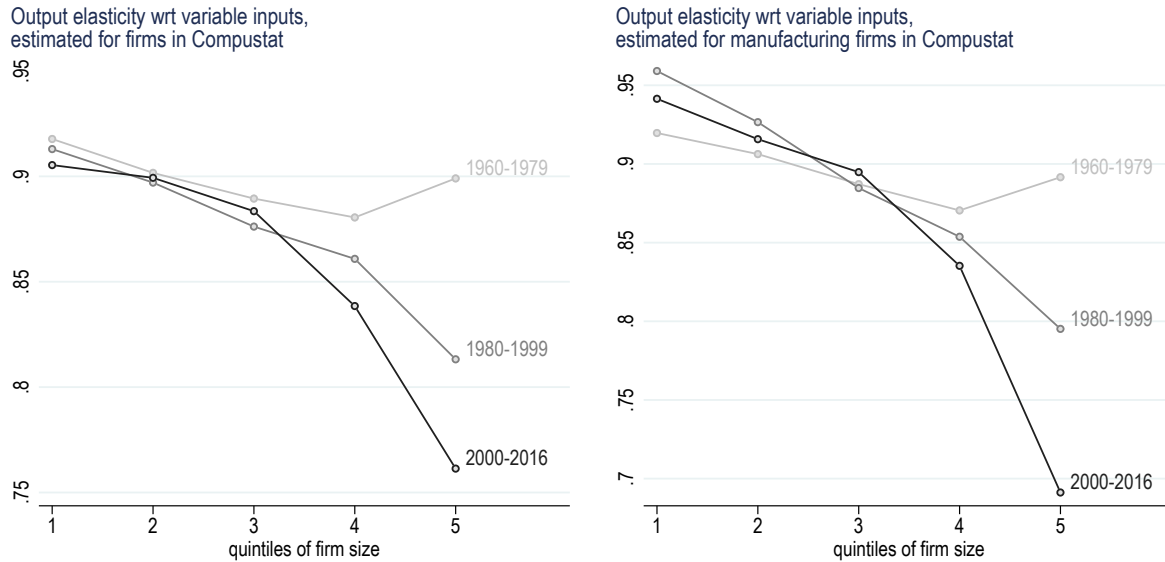


FIGURE 7: OUTPUT ELASTICITY WITH RESPECT TO VARIABLE INPUTS. The left panel presents estimates obtained for all Compustat firms. The right panel presents estimates obtained for Compustat firms in manufacturing.

Figure 7 summarizes our findings for all Compustat firms and for manufacturing. In the 60s and 70s, firms had similar variable input elasticities. However, from there on, we estimate a pronounced decline in the variable input elasticity that concentrates among the firms in the top quintiles of the sales distribution. This clockwise rotation is precisely what our model predicts and coincides with the motivating evidence discussed in the introduction and which pointed to the uneven adoption of capital-intensive technologies by large firms. Appendix F provides alternative estimates assuming that: i. h and g are given by cubic polynomials; ii. there are no ex-post shocks ϵ (so that no proxy variables

are needed and we can treat y_{ft} as true revenue); or iii. assuming that z_{ft} follows a linear Markov process, so that we can conduct the estimation using a dynamic panel approach. All these sets of alternative assumptions deliver similar results.

We recognize that estimating markups at the firm level relies on strong assumptions. Moreover, we only conduct this exercise for firms in Compustat which are not a representative sample of the U.S. economy. Our estimates must be interpreted with the same caution required to approach previous empirical estimates of markups relying on these data and on similar methods.

3.2 Markups and comparison to previous papers

Figure 8 plots the implied time series for markups using our estimates of output elasticities described in the previous section. The black line provides the aggregate markup, which we compute as a sales-weighted harmonic mean of firm-level markups. As discussed above, this is the relevant notion of an aggregate markup for the behavior of the aggregate labor share in an industry or the economy, as well as for welfare (Edmond, Midrigan and Xu, 2018; Baqaee and Farhi, 2020b). Our estimates for markups suggest that they have been quite stable over time, fluctuating around 1.2.

For comparison, the red line in the figure plots the aggregate markup that would result if we grouped firms by industry only and assumed that all firms in the same industry operated technologies with the same capital intensity. This series reveals a mild secular increase in the aggregate markup from 1.25 in 1960 and 1.2 in 1980 to 1.3 in recent years, which is broadly in agreement with the harmonic mean (or cost weighted) estimates in Edmond, Midrigan and Xu (2018). Finally, the dashed red line reports estimates of an *arithmetic* mean of sales-weighted markups obtained under the assumption that all firms in a given industry operate technologies with the same capital intensity. These estimates are in line with those reported by De Loecker, Eeckhout and Unger (2020), but as we will see, are not relevant for understanding the contribution of markups to the decline in the labor share.³²

³²Appendix F shows that the arithmetic mean of sales-weighted markups computed under the assumption that all firms in an industry operate the same technology will spuriously increase when some firms adopt technologies with different capital intensities. The harmonic mean does not suffer from this mechanical bias, providing an additional reason for looking at the harmonic mean rather than the arithmetic sales-weighted mean of firm markups, especially if one believes that firms are increasingly diverging in their use of technology, as is the case in the data.

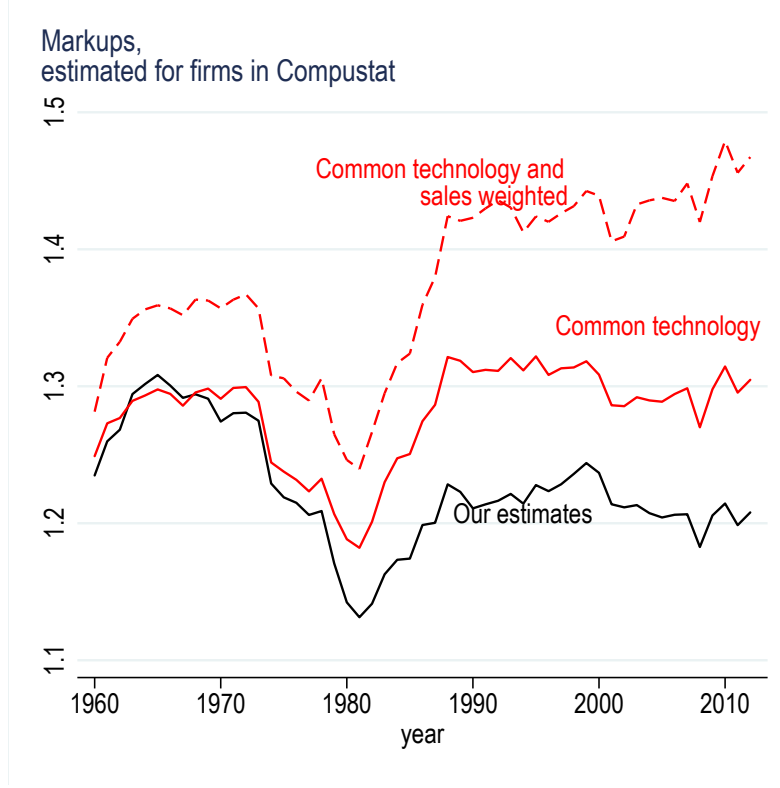


FIGURE 8: EVOLUTION OF MARKUPS. The figure presents the aggregate markup estimated for firms in Compustat. The black line provides our baseline estimates. The red line provides estimates obtained under the assumption of common revenue elasticities across firms in the same industry. The red dotted line provides a version of the later estimates where the aggregate markup is computed as the arithmetic sales weighted mean of firm-level markups.

3.3 Contribution to the decline in the labor share

As discussed above, the labor share in an industry at time t can be written as

$$s_{\ell t} := \frac{\varepsilon_{\ell t}}{\mu_t},$$

where $\varepsilon_{\ell t}$ denotes the share of labor in total industry costs, and μ_t is the industry markup, defined as a sales-weighted harmonic mean of firms' markups:

$$\frac{1}{\mu_t} := \sum_f \omega_{ft} \cdot \frac{1}{\mu_{ft}}.$$

The effects of technology, changes in factor prices, and substitution across firms with different factor intensities is captured by the cost share of labor in the industry, $\varepsilon_{\ell t}$. In what follows, we investigate the contribution of changes in markups and reallocation across firms with different markups to the labor share by looking at the percent change

over time in the inverse of the aggregate markup $1/\mu_t$. We focus on the period from 1980 to 2012, which matches the period studied in our model, and provide a number of extensions and robustness checks in Appendix F.

Figure 9 plots the cumulative percent change in $1/\mu_t$ between 1960 and each year aggregated across all industries. The figure presents our results separately for manufacturing firms and for non-manufacturing. Furthermore, following the exercises in Section 2.4, we decompose the contribution of markups into within-firm changes and allocative changes, which capture the contribution of reallocation towards firms with higher markups over time.³³ In line with what one would expect from an increase in competition since 1980, the reallocation component generates a negative contribution to the labor share, while within firm changes in markups generate a positive contribution.

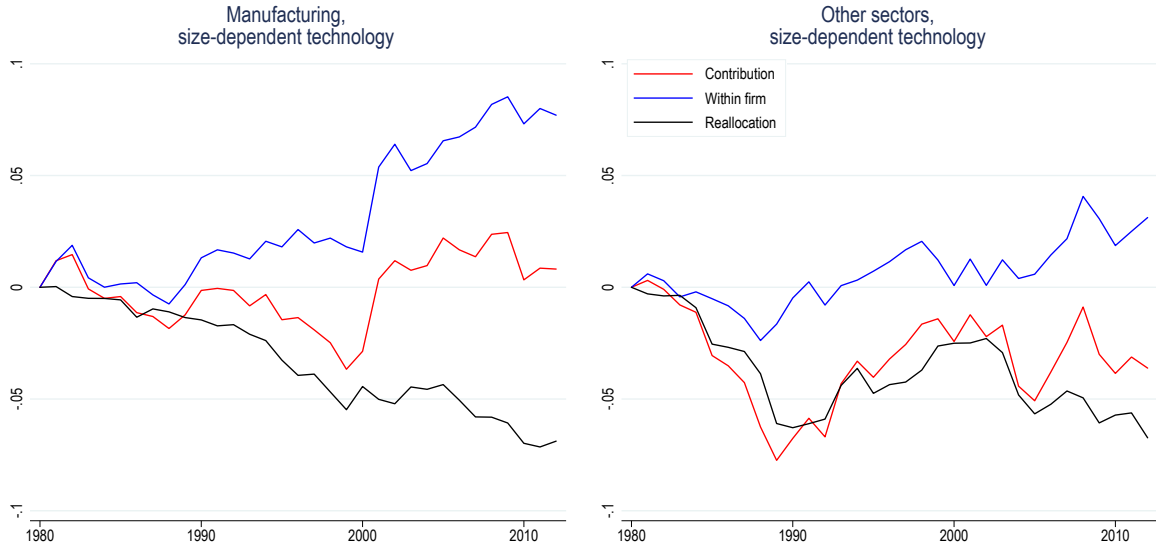


FIGURE 9: DECOMPOSITION OF THE CONTRIBUTION OF WITHIN-FIRM CHANGES IN MARKUPS AND BETWEEN-FIRM REALLOCATION TO (PERCENT) CHANGES IN THE LABOR SHARE. See the main text for details on this decomposition. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other economic sectors.

The patterns in Figure 9 suggest that the reallocation towards high-markup firms played a minor role in explaining the decline of the manufacturing labor share. In this sector, the reallocation component accounts for a decline of the labor share of 7% between 1980 and 2012—a quarter of the observed decline. These estimates provide an upper bound on the contribution of rising competition to the decline in the labor share of the

³³In particular, we compute the within-firm contribution to the percent change in markups in each year as $-\sum_f \omega_f \Delta \ln \mu_f$ and the reallocation component as $-\ln \mu_t + \sum_f \omega_f \Delta \ln \mu_f$. We then compute the cumulative contribution of within firm changes and the reallocation component over time.

sector, since we would expect the reallocation to higher markup firms induced by rising competition to be accompanied by a decrease in firm markups. Indeed, once we factor in the observed changes in firm markups, we find that the total contribution from markups to the labor share in manufacturing has been essentially zero.

Outside of manufacturing, we find that the reallocation component explained up to 6% of the decline in the labor share since 1980—about one half of the observed decline. In these sectors, the within component has been much weaker and close to zero. As a whole, the estimates support the idea that in these sectors, rising competition might have reduced the labor share via allocative changes towards high markup firms without bringing a similar offsetting reduction in firm markups.

As a whole, we view this exercise as supporting the quantitative conclusions derived in our model-based decomposition. In particular, the estimates here point to a crucial role for technology in the form of increased capital–labor substitution in explaining the decline in the labor share in manufacturing. Although more noisy, the estimates also suggest that rising competition and allocative changes might explain up to one half of the decline in the labor share outside of manufacturing.

This decomposition also highlights the importance of accounting for differences in technology across firms. If we counterfactually assumed that output elasticities are common to all firms in a given industry, as is commonly done in the literature, we would estimate a larger negative contribution of markups to the labor share, especially in manufacturing. Accounting for differences in technology modifies these conclusions and shows that previous estimates confound differences in the factor intensity of technology across firms in a given industry with differences in markups.³⁴

³⁴ An alternative approach to estimating markups assumes constant returns to scale (as we do) but tries to directly measure the user cost of capital as

$$R = r + \delta - \pi_k,$$

where r is a required rate of return inclusive of an industry-specific risk premium, δ is the depreciation rate, and π_k is the expected change over time in capital prices. One can then compute markups as revenue divided by total cost ($= V + RK$). The user-cost formula, which goes back to Hall and Jorgenson (1967) requires common and frictionless capital markets and assumes no adjustment costs for capital. In other words, this formula assumes that at all points in time, the marginal product of capital is equalized across all firms in a given industry. This strikes us as restrictive when thinking about firms undergoing a costly automation process. Instead, the approach used above makes no assumption about the marginal product of capital across firms, or the importance of adjustment costs. In any case, Baqaee and Farhi (2020b) report that, when using the user-cost formula above, the contribution of the between component to the decline of the labor share during 1985–2015 is 5%, about half of the observed decline.

4 CONCLUDING REMARKS

This paper starts from the observation that the adoption of modern automation technologies concentrates at large firms. Our main point is that, once we account for this vast heterogeneity in technology adoption, one can make sense of the dynamics of the labor share decline across firms both qualitatively and quantitatively. Moreover, accounting for these differences in technology adoption leads to different conclusions regarding the relative contribution of markups vs. technology to the decline in the labor share across industries.

We made this point in three related exercises:

1. First, we developed a model of firm dynamics with costly automation decisions to study the dynamics of labor shares, market shares, and capital-labor substitution across firms and industries. We first show that a special case of our model with a simple demand side—a standard CES aggregator that implies constant markups—produces firm-level labor share dynamics in response to falling capital prices that are qualitatively and quantitatively in line with the observed firm-level data in the U.S. manufacturing sector. In particular, we find that the model reproduces the striking fact that while the sectoral labor share declined drastically, the labor share of the median firm increased slightly. The model also explains a range of other related observations, ranging from specific labor share decompositions proposed in the literature to the observed increases in concentration and productivity dispersion.
2. Second, we extended our model to allow for a more general demand side—a log-concave demand system that implies markups which are increasing in firm size and productivity—, which allows us to account for the effects of rising competition and the ensuing reallocation towards larger high-markup firms. We used this model to quantitatively decompose the fall in the labor share and the rise in sales concentration into a component driven by lower capital prices and a component driven by rising competition. We find that the main drivers of the decline in the labor share vary by sector. In manufacturing, capital-labor substitution driven by falling capital prices accounts for the majority of the sectoral labor share decline. Instead, we find that the rise in competition is more important in other sectors, in particular in retail, where it accounts for up to 60% of the falling labor share.
3. Third, we empirically estimated the contribution of a more general form of reallocation to high-markup firms, which are not necessarily large in the data. This

extra flexibility comes at the expense of requiring firm-level estimates of markups and output elasticities. Using standard techniques to estimate production functions but allowing for technology to vary across firms of different size, we estimate that the variable input elasticity of large firms has declined over time. While this finding is also in line with the costly automation mechanism at the core of our paper, the main implication here is that accounting for technological heterogeneity substantially weakens the rise in the aggregate markup. This exercise largely confirms our model-based findings: reallocation to high-markup firms can explain only a small fraction of the labor share decline in manufacturing, while playing a significant role in other sectors.

Our paper motivates several avenues for future research. On the empirical front, we need more direct evidence on the causes and consequences of the heterogeneous adoption of modern capital-intensive technologies. The new technology modules in the US Census *Annual Business Survey* provide a promising tool for studying these questions. Moreover, our model points to the importance of developing estimates for markups and production function that can account in a flexible way for heterogeneity in technology, automation, and factor intensity across firms.

On the theory side, we need more work to understand the root causes of rising competition as well as more flexible quantitative models to account for more general specifications of markups and demand. Although the non-CES demand systems used here are gaining traction in macroeconomics, they only allow markups and passthroughs to be functions of firm size. Finally, in our work all firms are ex-ante equal and have the same efficiency at using capital. However, it would be interesting to explore the implications of allowing for permanent differences in capital efficiency in firm dynamics models, and the response of the economy to lower capital prices.

REFERENCES

- ACEMOGLU, D., C. LELARGE, AND P. RESTREPO (2020): “Competing with Robots: Firm-Level Evidence from France,” in *AEA Papers and Proceedings*, vol. 110, pp. 383–88.
- ACEMOGLU, D., AND P. RESTREPO (2018): “The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment,” *American Economic Review*, 108(6), 1488–1542.
- (2020): “Robots and jobs: Evidence from US labor markets,” *Journal of Political Economy*, 128(6), 2188–2244.
- ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2015): “Identification Properties of Recent Production Function Estimators,” *Econometrica*, 83(6), 2411–2451.

- AKERMAN, A., E. LEUVEN, AND M. MOGSTAD (2021): “Information Frictions, Internet, and the Relationship between Distance and Trade,” *American Economic Journal: Applied Economics*, (forthcoming).
- AMITI, M., O. ITSKHOKI, AND J. KONINGS (2019): “International shocks, variable markups, and domestic prices,” *The Review of Economic Studies*, 86(6), 2356–2402.
- ARKOLAKIS, C., A. COSTINOT, D. DONALDSON, AND A. RODRÍGUEZ-CLARE (2018): “The Elusive Pro-Competitive Effects of Trade,” *The Review of Economic Studies*, 86(1), 46–80.
- AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. VAN REENEN (2020): “The fall of the labor share and the rise of superstar firms,” *The Quarterly Journal of Economics*, 135(2), 645–709.
- BAQAEE, D., AND E. FARHI (2020a): “The Darwinian Returns to Scale,” Working Paper 27139, National Bureau of Economic Research.
- BAQAEE, D. R., AND E. FARHI (2020b): “Productivity and misallocation in general equilibrium,” *The Quarterly Journal of Economics*, 135(1), 105–163.
- BARKAI, S. (2020): “Declining labor and capital shares,” *The Journal of Finance*, LXXV(5).
- BERGER, D., K. HERKENHOFF, AND S. MONGEY (2019): “Labor Market Power,” Working Paper 25719, National Bureau of Economic Research.
- BOND, S., A. HASHEMI, G. KAPLAN, AND P. ZOCH (2020): “Some Unpleasant Markup Arithmetic: Production Function Elasticities and their Estimation from Production Data,” Discussion paper, University of Chicago.
- BONFIGLIOLI, A., R. CRINÒ, H. FADINGER, AND G. GANCIA (2020): “Robot Imports and Firm-Level Outcomes,” Discussion paper, CEPR Discussion Papers No. 14593.
- BRIDGMAN, B. (2018): “Is labor’s loss capital’s gain? Gross versus net labor shares,” *Macroeconomic Dynamics*, 22(8), 2070–2087.
- CLEMENTI, G. L., AND B. PALAZZO (2016): “Entry, Exit, Firm Dynamics, and Aggregate Fluctuations,” *American Economic Journal: Macroeconomics*, 8(3), 1–41.
- CORRADO, C., C. HULTEN, AND D. SICHEL (2009): “Intangible Capital and U.S. Economic Growth,” *Review of Income and Wealth*, 55(3), 661–685.
- CUMMINS, J. G., AND G. L. VIOLANTE (2002): “Investment-Specific Technical Change in the US (1947–2000): Measurement and Macroeconomic Consequences,” *Review of Economic Dynamics*, 5(2), 243–284.
- DAO, M. C., M. DAS, AND Z. KOCZAN (2019): “Why is labour receiving a smaller share of global income?,” *Economic Policy*, 34(100), 723–759.
- DE LOECKER, J., J. EECKHOUT, AND G. UNGER (2020): “The rise of market power and the macroeconomic implications,” *The Quarterly Journal of Economics*, 135(2), 561–644.
- DECKER, R. A., J. HALTIWANGER, R. S. JARMIN, AND J. MIRANDA (2020): “Changing Business Dynamism and Productivity: Shocks versus Responsiveness,” *American Economic Review*, (forthcoming).
- DICECIO, R. (2009): “Sticky wages and sectoral labor comovement,” *Journal of Economic Dynamics and Control*, 33(3), 538–553.
- DINLERSOZ, E. M., AND Z. WOLF (2018): “Automation, Labor Share, and Productivity: Plant-Level Evidence from U.S. Manufacturing,” Discussion paper, Center for Economic Studies Working Paper CES-WP-18-39.
- EDEN, M., AND P. GAGGL (2018): “On the Welfare Implications of Automation,” *Review of Economic Dynamics*, 29, 15–43.
- EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2018): “How costly are markups?,” Discussion paper, National Bureau of Economic Research.
- ELSBY, M. W., B. HOBIJN, AND A. ŞAHİN (2013): “The decline of the US labor share,”

- Brookings Papers on Economic Activity*, 2013(2), 1–63.
- FOSTER, L., J. HALTIWANGER, AND C. SYVERSON (2008): “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?,” *American Economic Review*, 98(1), 394–425.
- GORDON, R. J. (1990): *The Measurement of Durable Goods Prices*. University of Chicago Press, Chicago, IL.
- GUTIERREZ, G., AND S. PITON (2020): “Revisiting the Global Decline of the (Non-Housing) Labor Share,” *American Economic Review: Insights*, (forthcoming).
- HALL, R. E., AND D. W. JORGENSEN (1967): “Tax Policy and Investment Behavior,” *The American Economic Review*, 57(3), 391–414.
- HOPENHAYN, H. A. (1992): “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica*, 60(5), 1127–1150.
- HOUTHAKKER, H. S. (1955): “The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis,” *Review of Economic Studies*, 23(1), 27–31.
- HUBMER, J. (2020): “The Race Between Preferences and Technology,” Unpublished manuscript, University of Pennsylvania.
- HUMLUM, A. (2019): “Robot Adoption and Labor Market Dynamics,” Discussion paper, Mimeo, Princeton University.
- KARABARBOUNIS, L., AND B. NEIMAN (2013): “The Global Decline of the Labor Share*,” *The Quarterly Journal of Economics*, 129(1), 61–103.
- KAYMAK, B., AND I. SCHOTT (2018): “Corporate Tax Cuts and the Decline of the Labor Share,” 2018 Meeting Papers 943, Society for Economic Dynamics.
- KEHRIG, M., AND N. VINCENT (2020): “The Micro-Level Anatomy of the Labor Share Decline,” NBER Working Papers 25275, National Bureau of Economic Research, Inc.
- KIMBALL, M. S. (1995): “The Quantitative Analytics of the Basic Neomonetarist Model,” *Journal of Money, Credit and Banking*, 27(4), 1241–1277.
- KLENOW, P. J., AND J. L. WILLIS (2016): “Real Rigidities and Nominal Price Changes,” *Economica*, 83(331), 443–472.
- KOCH, M., I. MANUYLOV, AND M. SMOLKA (2019): “Robots and Firms,” Discussion paper, CESifo Working Paper No. 7608.
- KOH, D., R. SANTAELULIA-LLOPIS, AND Y. ZHENG (2020): “Labor share decline and intellectual property products capital,” *Econometrica*, (forthcoming).
- LASHKARI, D., A. BAUER, AND J. BOUSSARD (2019): “Information Technology and Returns to Scale,” Discussion paper, Boston College.
- LEE, Y., AND T. MUKOYAMA (2015): “Entry and exit of manufacturing plants over the business cycle,” *European Economic Review*, 77, 20 – 27.
- MELITZ, M. J., AND G. I. OTTAVIANO (2008): “Market size, trade, and productivity,” *The review of economic studies*, 75(1), 295–316.
- OBERFIELD, E., AND D. RAVAL (2014): “Micro Data and Macro Technology,” NBER Working Papers 20452, National Bureau of Economic Research, Inc.
- OLLEY, G. S., AND A. PAKES (1996): “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 64(6), 1263–1297.
- PERLA, J., C. TONETTI, AND M. E. WAUGH (2021): “Equilibrium Technology Diffusion, Trade, and Growth,” *American Economic Review*, 111(1), 73–128.
- PIKETTY, T. (2014): *Capital in the Twenty-First Century*. Harvard University Press, Cambridge.
- ROGNLIE, M. (2015): “Deciphering the Fall and Rise in the Net Capital Share: Accumulation

- or Scarcity?,” *Brookings Papers on Economic Activity*, 46(1 (Spring), 1–69.
- SMITH, M., D. YAGAN, O. ZIDAR, AND E. ZWICK (2019): “The Rise of Pass-Throughs and the Decline of the Labor Share,” Discussion paper, Princeton University.
- UNGER, G. (2019): “Scale-Biased Technological Change,” Discussion paper, Harvard University.
- ZEIRA, J. (1998): “Workers, Machines, and Economic Growth,” *The Quarterly Journal of Economics*, 113(4), 1091–1117.
- ZOLAS, N., Z. KROFF, E. BRYNJOLFSSON, K. MCELHERAN, D. BEEDE, C. BUFFINGTON, N. GOLDSCHLAG, L. FOSTER, AND E. DINLERSOZ (2020): “Measuring Technology Adoption in Enterprise-Level Surveys: The Annual Business Survey,” Discussion paper, US Census.

Online Appendix to “Not a Typical Firm: The Joint Dynamics of Firms, Labor Shares, and Capital–Labor Substitution”

Joachim Hubmer and Pascual Restrepo

March 11, 2021

A PROOFS FOR THE CES-DEMAND MODEL

A.1 Lemmas and propositions in the main text

This section provides proofs for Lemma 1 and Propositions 1–3 in the main text.

Proof of Lemma 1. We first prove that optimal firm choices are given by $\alpha' = \max\{\hat{\alpha}_t(z), \alpha\}$. Consider a firm that sets α' above α at time t . The first-order condition pinning down this choice is given by

$$c_a \cdot y_t = \frac{1}{1+r} \mathbb{E} \left[\frac{\partial \pi_{t+1}(\alpha', z_{t+1})}{\partial \alpha'} + c_a \cdot y_{t+1} \middle| z_t \right].$$

We can rewrite this as

$$c_a \cdot ((1+r)y_t - y_{t+1}) = \mathbb{E} \left[\sigma^{-\sigma} (\sigma - 1)^{\sigma-1} \cdot y_{t+1} \cdot z_{t+1}^{\sigma-1} \cdot \frac{\partial \tilde{c}_{t+1}(\alpha')^{1-\sigma}}{\partial \alpha'} \middle| z_t \right].$$

By assumption, the left hand side is a positive constant. Moreover, the right-hand side is increasing in z_t and decreasing in α' . It follows that the first-order condition implicitly defines the optimal choice of α' as an increasing function of z_t , $\hat{\alpha}_t(z)$. The restriction that $\alpha' \geq \alpha$ then implies that $\alpha' = \max\{\hat{\alpha}_t(z), \alpha\}$ as wanted.

We now turn to the limiting properties of $\hat{\alpha}_t(z)$. Since z follows an AR1 in logs, given $\rho_z > 0$ it holds that $\lim_{z_t \rightarrow 0} \mathbb{E}[z_{t+1}^{\sigma-1}] = 0$. Thus, the period $t+1$ benefit of automation converges to zero as z_t converges to zero. For any fixed cost $c_a > 0$, automating a positive measure of tasks $\Delta\alpha > 0$ in period t cannot be optimal, as doing so in $t+1$ would reduce the present discounted value of automation costs by a discrete amount: $\frac{c_a y_{t+1} \Delta\alpha}{1+r} < c_a y_t \Delta\alpha$, which holds given the assumption that aggregate output y_t grows at a rate below r . Likewise, since $\lim_{z_t \rightarrow \infty} \mathbb{E}[z_{t+1}^{\sigma-1}] = \infty$, for any $\alpha < \alpha_{t+1}^*$ the marginal automation benefit exceeds the (finite) marginal automation cost for large enough z_t . ■

Proof of Proposition 1. Suppose that $q_t(x) = q(x)$. Let's consider an allocation in which $\alpha = \alpha^*$ for all firms. Incumbent firms set $\alpha' = \max\{\hat{\alpha}_t(z), \alpha\}$. Because $\hat{\alpha}_t(z) \leq \alpha^*$,

incumbents keep $\alpha' = \alpha^*$. Finally, entrants start with $\alpha = \alpha_{0t} = \alpha^*$, which shows that this allocation is a stationary equilibrium. ■

Note: because of the fixed cost structure, the economy admits multiple stationary equilibria. In particular, an allocation with all firms having a common automation level $\alpha \in [\alpha^*, 1]$ also corresponds to an equilibrium. However, if we start from an allocation with $\alpha \leq \alpha^*$ for all firms, the economy will converge to the stationary equilibrium with $\alpha = \alpha^*$ for all firms.

Proof of Proposition 2. To account for a proportional increase in $q(x)$, let's write $q(x) = q \cdot q_0(x)$. Consider a steady state, where $\alpha = \alpha^*$ for all firms. We are interested in the comparative statics of the stationary equilibrium with aggregate equilibrium objects (w, y, α^*) as q changes by $d \ln q$.

A firm's cost function can be written as

$$c(z; w, q, \alpha^*) = \frac{1}{z} \cdot \tilde{c}(w, q, \alpha^*),$$

where

$$\tilde{c}(w, q, \alpha^*) = \left(\Psi_k(\alpha^*) \cdot q^{\eta-1} + \Psi_\ell(\alpha^*) \cdot w^{1-\eta} \right)^{\frac{1}{1-\eta}}$$

is the unit cost function of a firm with unitary productivity.

We claim that \tilde{c} is unchanged across steady states. To see this, observe that if the claim is true, then firm profits are proportional to y : $\pi(z; y) \propto z^{\sigma-1} \cdot y$. Since fixed costs are also scaled by y , the value function is linear in y , and across steady states entry and exit decisions as a function of z are unchanged. Consequently, the distribution of firm productivities with pdf $f(z)$ does not change across steady states. We now turn to the ideal-price index condition in equation (3), which can be written in a stationary equilibrium as

$$\int_z \mu^{1-\sigma} \cdot z^{\sigma-1} \cdot \tilde{c}(w, q, \alpha^*)^{1-\sigma} \cdot f(z) \cdot dz = 1,$$

and implies that the unit cost $\tilde{c}(w, q, \alpha^*)$ is constant, verifying our claim.

An application of Shephard's lemma implies that

$$d \ln \tilde{c} = \varepsilon_\ell \cdot d \ln w - \varepsilon_k \cdot d \ln q.$$

The envelope theorem ensures that the effect of changes in α^* on \tilde{c} are second order and

can be ignored. Because $d \ln \tilde{c} = 0$, we can solve for the change in wages as

$$d \ln w = \frac{\varepsilon_k}{\varepsilon_\ell} \cdot d \ln q = \frac{1 - \varepsilon_\ell}{\varepsilon_\ell} \cdot d \ln q > 0.$$

We now turn to the behavior of cost shares. Along the stationary equilibrium, all firms have the same labor cost share, which is given by

$$\varepsilon_\ell = \frac{\Psi_\ell(\alpha^*) \cdot w^{1-\eta}}{\Psi_k(\alpha^*) \cdot q^{\eta-1} + \Psi_\ell(\alpha^*) \cdot w^{1-\eta}}.$$

This common cost share for labor will vary with prices and α^* . Equation (5) implies that the change in the optimal threshold α^* satisfies

$$d \ln \alpha^* = \frac{1}{\partial \ln \psi_\ell(\alpha^*) / (q(\alpha^*) \cdot \psi_k(\alpha^*)) / \partial \ln \alpha} (d \ln q + d \ln w).$$

Using this expression for $d \ln \alpha^*$ and the definition of η_{LR} , we can compute the change in the cost share of labor as

$$\begin{aligned} d \ln \varepsilon_\ell &= (1 - \varepsilon_\ell) d \ln \frac{\varepsilon_\ell}{\varepsilon_k} \\ &= (1 - \varepsilon_\ell)(1 - \eta)(d \ln w + d \ln q) + (1 - \varepsilon_\ell) \frac{\partial \ln \Psi_\ell(\alpha^*) / \Psi_k(\alpha^*)}{\partial \ln \alpha} d \ln \alpha^* \\ &= (1 - \varepsilon_\ell)(1 - \eta)(d \ln w + d \ln q) + (1 - \varepsilon_\ell) \frac{\partial \ln \Psi_\ell(\alpha^*) / \Psi_k(\alpha^*) / \partial \ln \alpha}{\partial \ln \psi_\ell(\alpha^*) / (q(\alpha^*) \cdot \psi_k(\alpha^*)) / \partial \ln \alpha} (d \ln q + d \ln w) \\ &= (1 - \varepsilon_\ell)(1 - \eta)(d \ln w + d \ln q) + (1 - \varepsilon_\ell)(\eta - \eta_{LR})(d \ln q + d \ln w) \\ &= (1 - \varepsilon_\ell)(1 - \eta_{LR})(d \ln q + d \ln w), \end{aligned}$$

which using the formula above for the change in wages can be written as

$$d \ln \varepsilon_\ell = \frac{1 - \varepsilon_\ell}{\varepsilon_\ell} (1 - \eta_{LR}) d \ln q.$$

Finally, for a firm that is small enough such that it does not automate along the transition, the capital–labor elasticity is equal to the task-level substitution elasticity η . For this firm, we have

$$d \ln \varepsilon_\ell(\theta) = (1 - \varepsilon_\ell(\theta)) d \ln \frac{\varepsilon_\ell(\theta)}{\varepsilon_k(\theta)} = (1 - \varepsilon_\ell)(1 - \eta)(d \ln w + d \ln q),$$

which using the formula above for the change in wages can be written as

$$d \ln \varepsilon_\ell(\theta) = \frac{1 - \varepsilon_\ell(\theta)}{\varepsilon_\ell(\theta)} (1 - \eta) d \ln q.$$

■

Proof of Proposition 3. The reasoning is analogous to 2. To account for a proportional increase in $q(x)$ for $x > \alpha^*$, let's write $q(x) = q \cdot q_0(x)$ for $x > \alpha^*$ and $q(x) = q_0(x)$ otherwise. We are interested in the comparative statics of the stationary equilibrium with aggregate equilibrium objects (w, y, α^*) as q changes from 1 by $d \ln q$.

First, recall that $\tilde{c}(w, q, \alpha^*)$ is the minimum cost of production given w and q . An increase in q thus reduces \tilde{c} once we account for changes in α^* , which implies that w increases. Thus, we have $d \ln w > 0$. Note that the first-order approximation used in the proof of Proposition 2 yields $d \ln w = 0$. This is because the increase in wages is second order but positive nonetheless.

We now turn to the behavior of cost shares. Along the stationary equilibrium, all firms have the same labor cost share, which is given by

$$\varepsilon_\ell = \frac{\Psi_\ell(\alpha^*) \cdot w^{1-\eta}}{\Psi_k(\alpha^*) + \Psi_\ell(\alpha^*) \cdot w^{1-\eta}}.$$

This common cost share for labor will vary with prices and α^* . Equation (5) implies that the change in the optimal threshold α^* satisfies

$$d \ln \alpha^* = \frac{1}{\partial \ln \psi_\ell(\alpha^*) / (q(\alpha^*) \cdot \psi_k(\alpha^*)) / \partial \ln \alpha} (d \ln q + d \ln w).$$

Using this expression for $d \ln \alpha^*$ and the definition of η_{LR} , we can compute the change in the cost share of labor as

$$\begin{aligned} d \ln \varepsilon_\ell &= (1 - \varepsilon_\ell) d \ln \frac{\varepsilon_\ell}{\varepsilon_k} \\ &= (1 - \varepsilon_\ell)(1 - \eta) d \ln w + (1 - \varepsilon_\ell) \frac{\partial \ln \Psi_\ell(\alpha^*) / \Psi_k(\alpha^*)}{\partial \ln \alpha} d \ln \alpha^* \\ &= (1 - \varepsilon_\ell)(1 - \eta) d \ln w + (1 - \varepsilon_\ell) \frac{\partial \ln \Psi_\ell(\alpha^*) / \Psi_k(\alpha^*) / \partial \ln \alpha}{\partial \ln \psi_\ell(\alpha^*) / (q(\alpha^*) \cdot \psi_k(\alpha^*)) / \partial \ln \alpha} (d \ln q + d \ln w) \\ &= (1 - \varepsilon_\ell)(1 - \eta) d \ln w + (1 - \varepsilon_\ell)(\eta - \eta_{LR})(d \ln q + d \ln w) \\ &= -(1 - \varepsilon_\ell)(\eta_{LR} - \eta) d \ln q + (1 - \varepsilon_\ell)(1 - \eta_{LR}) d \ln w. \end{aligned}$$

Finally, for a firm that is small enough such that it does not automate along the

transition, the capital–labor elasticity is equal to the task-level substitution elasticity η . For this firm, we have the same expression derived above:

$$d \ln \varepsilon_\ell(\theta) = (1 - \varepsilon_\ell(\theta))(1 - \eta)(d \ln w + d \ln q).$$

■

A.2 Additional results and formulas

This section derives the formula for the long-run elasticity of substitution provided in the main text. It also derives the approximation for the productivity gains from automation given in footnote 18.

Long-run K-L elasticity of substitution η_{LR} : As usual, we define this elasticity as the partial equilibrium response of the capital–labor ratio to a change in the relative price $\tilde{w} \equiv \frac{w}{1/q}$; here taking into account the change in α^* . This long-run elasticity of substitution is therefore defined implicitly by the identity

$$d \ln \frac{\varepsilon_\ell}{\varepsilon_k} \equiv (1 - \eta_{LR}) \cdot d \ln \tilde{w}.$$

We have that

$$(8) \quad d \ln \frac{\varepsilon_\ell}{\varepsilon_k} = d \ln \frac{\Psi_\ell(\alpha) \cdot w^{1-\eta}}{\Psi_k(\alpha) \cdot q^{\eta-1}} = \frac{\partial \ln \Psi_\ell(\alpha) / \Psi_k(\alpha)}{\partial \ln \alpha} \cdot \frac{\partial \ln \alpha^*}{\partial \ln \tilde{w}} \cdot d \ln \tilde{w} + (1 - \eta) \cdot d \ln \tilde{w}.$$

Equation (5) implies that

$$\frac{\partial \ln \alpha^*}{\partial \ln \tilde{w}} = \left(\frac{\partial \ln \psi_\ell(\alpha) / (q(\alpha) \cdot \psi_k(\alpha))}{\partial \ln \alpha} \right)^{-1}.$$

Plugging this in equation (8) we obtain

$$d \ln \frac{\varepsilon_\ell}{\varepsilon_k} = \left[1 - \underbrace{\left(\eta + \frac{\partial \ln \Psi_k(\alpha) / \Psi_\ell(\alpha)}{\partial \ln \alpha} \right) \left(\frac{\partial \ln \psi_\ell(\alpha) / (q(\alpha) \cdot \psi_k(\alpha))}{\partial \ln \alpha} \right)}_{\equiv \eta_{LR}} \right] d \ln \tilde{w}$$

which gives the formula for η_{LR} provided in the main text.

Productivity gains from automation: Consider an increase in the automation threshold of $d\alpha > 0$ for a firm with $\alpha < \alpha^*$. The contribution of this technology to a firm's

TFP is then given by

$$d \ln \text{TFP}_\alpha := -d \ln c(\theta, w) = -\frac{1}{1-\eta} \left[\varepsilon_\ell(\alpha, w) \frac{\Psi'_k(\alpha)}{\Psi_k(\alpha)} + \varepsilon_\ell(\alpha, w) \frac{\Psi'_\ell(\alpha)}{\Psi_\ell(\alpha)} \right] d\alpha.$$

We can rewrite this expression as follows:

$$\begin{aligned} d \ln \text{TFP}_\alpha &= -\frac{1}{1-\eta} \left[\varepsilon_\ell(\alpha, w) \frac{(q(\alpha) \cdot \psi_k(\alpha))^{\eta-1}}{\Psi_k(\alpha)} - \varepsilon_\ell(\alpha, w) \frac{\psi_\ell(\alpha)^{\eta-1}}{\Psi_\ell(\alpha)} \right] \cdot d\alpha \\ &= -\frac{1}{1-\eta} \left[\frac{(q(\alpha) \cdot \psi_k(\alpha))^{\eta-1}}{\tilde{c}^{1-\eta}} - \frac{\psi_\ell(\alpha)^{\eta-1} w^{1-\eta}}{\tilde{c}^{1-\eta}} \right] \cdot d\alpha \geq 0. \end{aligned}$$

Log-linearizing the function $\frac{1}{1-\eta} x^{1-\eta}$, we can approximate the productivity gains when $\alpha \rightarrow \alpha^*$ as

$$d \ln \text{TFP}_\alpha \approx \left[\ln \left(\frac{w}{\psi_\ell(\alpha)} \right) - \ln \left(\frac{1}{q(\alpha) \cdot \psi_k(\alpha)} \right) \right] \cdot \frac{(w/\psi_\ell(\alpha))^{1-\eta}}{\tilde{c}^{1-\eta}} \cdot d\alpha \geq 0,$$

which yields the formula in the main text.

B PROOFS FOR THE MODEL WITH VARIABLE MARKUPS

B.1 Properties of the Klenow–Willis specification

As a convenient functional form for the Kimball (1995) aggregator $H(\cdot)$ we use the specification from Klenow and Willis (2016), defined as

$$H(\bar{y}) \equiv 1 + (\sigma - 1) \cdot \exp \left(\frac{1}{\nu} \right) \cdot \nu^{\frac{\sigma}{\nu}-1} \cdot \left[\Gamma \left(\frac{\sigma}{\nu}, \frac{1}{\nu} \right) - \Gamma \left(\frac{\sigma}{\nu}, \frac{\bar{y}^{\frac{\nu}{\sigma}}}{\nu} \right) \right],$$

where $\bar{y} = \frac{y(\theta)}{\lambda \cdot y}$ is the relative quantity of a variety, and $\Gamma(\cdot, \cdot)$ is the upper incomplete Gamma function,

$$\Gamma(s, x) \equiv \int_x^\infty t^{s-1} \cdot \exp(-t) dt.$$

This specification gives rise to the following (relative) demand function $D^{-1} = H'$:

$$D(\bar{p}) = \left(1 - \nu \cdot \ln\left(\bar{p} \cdot \frac{\sigma}{\sigma-1}\right)\right)^{\frac{\sigma}{\nu}},$$

$$D'(\bar{p}) = \frac{\sigma}{\bar{p}} \cdot \left(1 - \nu \cdot \ln\left(\bar{p} \cdot \frac{\sigma}{\sigma-1}\right)\right)^{\frac{\sigma}{\nu}-1},$$

where $\bar{p} = \frac{p(\theta)}{\rho}$ is the relative price of a variety. Thus, the price elasticity of demand is

$$(9) \quad -\frac{D'(\bar{p}) \cdot \bar{p}}{D(\bar{p})} = \frac{\sigma}{1 - \nu \cdot \ln\left(\bar{p} \cdot \frac{\sigma}{\sigma-1}\right)} = \sigma \cdot D(\bar{p})^{-\frac{\nu}{\sigma}},$$

which reduces to the constant σ if $\nu = 0$ (the benchmark case of a CES aggregator). In general, equation (9) shows that under this parametrization, the supra-elasticity of demand is equal to the constant $-\frac{\nu}{\sigma}$.

Static profit maximization: As usual, static optimal pricing relates the demand elasticity to the optimal markup $\mu = p/c$ as

$$\frac{\mu}{\mu-1} = -\frac{D'(\bar{p}) \cdot \bar{p}}{D(\bar{p})}.$$

Since $\bar{p} = \frac{p}{\rho} = \frac{\mu c}{\rho}$, we can re-write this equation using (6) such that it implicitly defines the optimal markup μ^* as a function of the relative cost $\hat{c} = c/\rho$,

$$\sigma \cdot \left(1 - \frac{1}{\mu^*(\hat{c})}\right) = 1 - \nu \cdot \ln\left(\frac{\sigma}{\sigma-1} \cdot \hat{c} \cdot \mu^*(\hat{c})\right).$$

Since the left-hand-side is an increasing function of $\mu^*(\hat{c})$ and the right-hand-side a decreasing function of $\hat{c} \cdot \mu^*(\hat{c})$, the optimal markup is implicitly defined as a strictly decreasing function of (relative) cost. Moreover, setting $\mu^* = 1$, we can solve for the (relative) choke price as

$$\bar{p}^{choke} = \frac{\sigma-1}{\sigma} \cdot \exp\left(\frac{1}{\nu}\right),$$

implying that firms with cost $c < \bar{p}^{choke} \cdot \rho$ do not produce.

Marshall's second laws: Equation (6) shows that the demand elasticity is increasing in the relative price and greater than 1 (Marshall's weak second law), imposing the restriction that $\sigma > 1$ and $\nu > 0$. To see that the strong law holds as well, write the logarithm of

marginal revenue as

$$\begin{aligned}\ln\left(\bar{p} + \frac{D(\bar{p})}{D'(\bar{p})}\right) &= \ln \bar{p} + \ln\left(1 + \frac{D(\bar{p})}{D'(\bar{p}) \cdot \bar{p}}\right) \\ &= \ln \bar{p} + \ln\left(\frac{\sigma + \nu \cdot \ln(\bar{p}) + \nu \cdot \ln\left(\frac{\sigma}{\sigma-1}\right) - 1}{\sigma}\right),\end{aligned}$$

which is a concave function of $\ln \bar{p}$ as desired.

B.2 Proofs and derivations

Before turning to the proofs of the propositions in the text, we derive the equations used in the characterization of the equilibrium. We will ignore time subscripts as long as it causes no confusion.

The demand for each variety is obtained by solving the following cost minimization problem:

$$\min_{y(\theta)} \int_{\theta} p(\theta) \cdot y(\theta) \cdot m(\theta) \cdot d\theta \quad \text{s.t.:} \quad \int_{\theta} \lambda \cdot H\left(\frac{y(\theta)}{\lambda \cdot y}\right) m(\theta) d\theta = 1.$$

Let $\rho \cdot y$ denote the Lagrange multiplier on the constraint. The first-order condition for the choice of $y(\theta)$ is then

$$p(\theta) = \rho \cdot H'\left(\frac{y(\theta)}{\lambda \cdot y}\right) \quad \Rightarrow \quad y(\theta) = y \cdot \lambda \cdot D\left(\frac{p(\theta)}{\rho}\right).$$

Moreover, because the price of the final good is normalized to 1, we must have

$$y = \int_{\theta} p(\theta) \cdot y(\theta) \cdot m(\theta) \cdot d\theta \quad \Rightarrow \quad 1 = \int_{\theta} \lambda \cdot p(\theta) \cdot D\left(\frac{p(\theta)}{\rho}\right) \cdot m(\theta) \cdot d\theta,$$

which is the ideal-price index condition given in the text.

Finally, plugging the demand for each variety in the constraint, we obtain

$$\int_{\theta} \lambda \cdot H\left(D\left(\frac{p(\theta)}{\rho}\right)\right) m(\theta) d\theta = 1,$$

which pins down the competitors' price index ρ .

The labor market clearing condition follows from an application of Roy's lemma, which implies that $c_w(\theta, w)$ yields the demand of labor per unit of goods sold.

Proof of Proposition 4. Prices are given by

$$p(c) = \arg \max_p y \cdot \lambda \cdot D\left(\frac{p}{\rho}\right) \cdot (p - c).$$

This problem has increasing differences in p and c , which implies that $p(c)$ is increasing in c .

Moreover, the first order condition for this problem is

$$-\frac{1}{\rho} D'\left(\frac{p}{\rho}\right) \cdot (p - c) = D\left(\frac{p}{\rho}\right) \Rightarrow \frac{\mu(c)}{\mu(c) - 1} = -\frac{p(c)}{\rho} \frac{D'\left(\frac{p(c)}{\rho}\right)}{D\left(\frac{p(c)}{\rho}\right)}.$$

Marshall's weak second law combined with the fact that $p(c)$ increases in c implies that the right-hand side of the above equation increases in c . The left-hand side is a decreasing function of $\mu(c)$, which therefore implies that $\mu(c)$ is decreasing in c as wanted.

We can rewrite the first-order condition for prices as

$$\frac{p(c)}{\rho} + \frac{D(p(c)/\rho)}{D'(p(c)/\rho)} = \frac{c}{\rho}.$$

Differentiating this expression yields

$$\frac{\partial \ln p(c)}{\partial \ln c} = 1 \Big/ n\left(\frac{p(c)}{\rho}\right),$$

where

$$n(x) = \frac{\partial \ln (x + D(x)/D'(x))}{\partial \ln x}$$

is a decreasing function according to Marshall's strong second law. It follows that $\ln p(c)$ is a convex function in $\ln c$ as wanted. Moreover, $\ln \mu(c) = \ln p(c) - \ln c$ will inherit this convexity.

Turning to sales, we have that $s(c)$ can be written as

$$s(c) = h(p(c)),$$

where $h(x) = xD(x)$ is a log-concave and decreasing function of x (from Marshall's weak second law). Thus, $s(c)$ is the composition of a log-concave and decreasing function ($h(x)$) with a log-convex and increasing function $p(c)$, which results in a log-concave and

decreasing function. ■

Proof of Proposition 5. In the next proofs, it will be convenient to define

$$\bar{c} := \frac{\tilde{c}}{\rho} \quad \bar{p} := \frac{p(c)}{\rho},$$

where recall that \tilde{c} is the unit cost for a firm with unitary productivity.

We can rewrite the firm problem as

$$\bar{p}\left(\frac{1}{z}\bar{c}\right) = \arg \max_{\bar{p}} D(\bar{p}) \cdot \left(\bar{p} - \frac{1}{z}\bar{c}\right).$$

Proposition 4 has three implications.

First, we can write prices as $p(z) = \rho \cdot \bar{p}\left(\frac{1}{z}\bar{c}\right)$, where $\bar{p}(\cdot)$ is an increasing function.

Second, the resulting markup $\mu(z)$ is given by

$$\mu(z) = \bar{\mu}\left(\frac{1}{z}\bar{c}\right),$$

where the function $\bar{\mu}(\cdot)$ is decreasing and log-convex.

Finally, sales are given by

$$s(z) = \bar{s}\left(\frac{1}{z}\bar{c}\right) = \lambda \cdot \rho \cdot \bar{p}\left(\frac{1}{z}\bar{c}\right) \cdot D\left(\bar{p}\left(\frac{1}{z}\bar{c}\right)\right),$$

where $\bar{s}(\cdot)$ is a decreasing and log-concave function.

The implicit definition of the competitors' price index implies

$$\int_z \lambda \cdot H\left(D\left(\bar{p}\left(\frac{1}{z}\bar{c}\right)\right)\right) f(z) dz = 1.$$

Thus, as λ increases, \bar{c} must increase. This implies that the effect of an increase in λ on prices, markups, and sales shares is isomorphic to that of an increase in \bar{c} .

We now turn to characterizing the effects of an increase in \bar{c} .

First, we have that for a given z , $\mu(z) = \bar{\mu}\left(\frac{1}{z}\bar{c}\right)$ will be decreasing in \bar{c} , as wanted.

Second, because the function $\bar{\mu}$ is log-convex, we have that, for $z > z'$,

$$\ln \mu(z) - \ln \mu(z') = \ln \bar{\mu}\left(\frac{1}{z}\bar{c}\right) - \ln \bar{\mu}\left(\frac{1}{z'}\bar{c}\right)$$

is decreasing in \bar{c} .

Third, because the function \bar{s} is log-concave, we have that, for $z > z'$,

$$\ln s(z) - \ln s(z') = \ln \bar{s}\left(\frac{1}{z}\bar{c}\right) - \ln \bar{s}\left(\frac{1}{z'}\bar{c}\right)$$

is increasing in \bar{c} . Note that there is a common λ term multiplying both $s(z)$ and $s(z')$ which cancels when computing the difference $\ln s(z) - \ln s(z')$. Thus, for relative shares, the effect of λ continues to be isomorphic to that of an increase in \bar{c} . ■

Proof of Proposition 6. As before, we investigate the implications of an increase in \bar{c} . We can write the aggregate markup as

$$\frac{1}{\mu} = \int_z \frac{1}{\bar{\mu}\left(\frac{1}{z}\bar{c}\right)} \cdot \bar{s}\left(\frac{1}{z}\bar{c}\right) \cdot f(z) \cdot dz \Bigg/ \int_z \bar{s}\left(\frac{1}{z}\bar{c}\right) \cdot f(z) \cdot dz.$$

With the change of variable $x = \frac{1}{z}\bar{c}$, we can rewrite this as

$$\frac{1}{\mu} = \int_x \frac{1}{\bar{\mu}(x)} \cdot g(x, \bar{c}) \cdot dx,$$

where $g(x, \bar{c})$ is a density function given by

$$g(x, \bar{c}) = \bar{s}(x) \cdot f(\bar{c}/x) \cdot dx \Bigg/ \int_x \bar{s}(x) \cdot f(\bar{c}/x) \cdot dx.$$

First, suppose that $f(z)$ is log-concave. This implies that

$$\ln \bar{s}(x) + \ln f(\bar{c}/x)$$

has increasing differences in x and \bar{c} . This is equivalent to the following monotone likelihood ratio property (MLRP):

$$\frac{g(x, \bar{c})}{g(x', \bar{c})} \text{ increasing in } \bar{c} \text{ for } x > x'.$$

The MLRP property implies that an increase in \bar{c} generates a shift up (in the first-order stochastic dominance sense) in $g(x, \bar{c})$. Because the function $\frac{1}{\bar{\mu}(x)}$ is increasing in x , the aggregate markup μ is decreasing in \bar{c} as wanted.

Second, suppose that $f(z)$ is log-convex. This implies that

$$\ln \bar{s}(x) + \ln f(\bar{c}/x)$$

has decreasing differences in x and \bar{c} . This is equivalent to the following monotone likelihood ratio property (MLRP):

$$\frac{g(x, \bar{c})}{g(x', \bar{c})} \text{ decreasing in } \bar{c} \text{ for } x > x'.$$

The MLRP property implies that an increase in \bar{c} generates a shift down (in the first-order stochastic dominance sense) in $g(x, \bar{c})$. Because the function $\frac{1}{\bar{\mu}(x)}$ is increasing in x , the aggregate markup μ is increasing in \bar{c} as wanted.

Finally, suppose that $f(z)$ is log-linear. This implies that

$$\ln \bar{s}(x) + \ln f(\bar{c}/x)$$

is a linear function in \bar{c} . Equivalently,

$$\frac{g(x, \bar{c})}{g(x', \bar{c})} \text{ independent of } \bar{c}.$$

Thus, the integral defining μ is independent of \bar{c} as wanted. ■

C ROBUSTNESS OF MODEL-BASED RESULTS

In this section, we discuss the robustness of our quantitative findings in Section 1 to the timing of automation decisions, as well as to different values for the short- and long-run capital–labor elasticity.

C.1 Timing of automation decisions

In the main text, we assume that firms invest in the adoption of new capital technologies (raising their α) in period t before the realization of their productivity in period $t + 1$, when the new technology first becomes operative. We have experimented with other timing assumptions and found that our results are robust on this dimension. Specifically, column (3) in Table 7 shows the calibration for an alternative model version where firms decide in the beginning of period $t + 1$, after their new productivity draw has materialized, whether to pay the operating fixed cost and whether to adopt new capital technologies, which are immediately operative. Column (3) in Table 8 shows the quantitative results over the transition period. We find that the results largely agree with our findings from the benchmark model, which are re-produced in column (2). As expected, since firms

with a high z (i.e., large firms) can raise their α immediately, the cross-dynamics term is somewhat larger in absolute value, going from -0.143 to -0.157. The inferred automation fixed cost c_a falls from 0.23 to 0.18, which reflects a lower option value of automation for small firms.

TABLE 7: Model robustness: Steady state calibration

		DATA	MODEL			
			BENCH- MARK	ALTERN. TIMING	LOWER η	$\eta_{LR} < 1$
		(1)	(2)	(3)	(4)	(5)
<i>I. Parameters</i>						
$\ln q_0$	Inverse capital price		-5.35	-5.36	-5.33	3.82
σ	Demand elasticity		11.0	11.0	11.0	11.0
σ_z	Std. dev. of $\ln z$ innovations		0.094	0.094	0.094	0.094
\underline{c}_f	Minimum fixed cost ($\times 10^{-6}$)		6.0	9.0	6.0	6.0
ξ	Dispersion fixed cost		0.330	0.350	0.330	0.330
μ_e	Entrant productivity		0.935	0.944	0.936	0.935
η	Task substitution elasticity		0.40	0.40	0.20	0.40
γ	Comparative advantage*		0.95	0.95	1.15	5.0*
<i>II. Moments</i>						
	Manufacturing labor share	60.1%	60.2%	60.1%	60.1%	60.1%
	Aggregate markup	1.10	1.10	1.10	1.10	1.10
	Top 4 firms' sales share	40.0%	40.1%	40.1%	40.1%	40.1%
	Entry (=exit) rate	0.062	0.063	0.063	0.063	0.063
	Size of exiters	0.490	0.491	0.490	0.491	0.491
	Size of entrants	0.600	0.598	0.600	0.602	0.598
	Short-run K-L elasticity		0.40	0.40	0.20	0.40
	Long-run K-L elasticity		1.35	1.35	1.35	0.90*

Notes: See Section 1 and Table 1 for the calibration of the benchmark model (corresponding to the manufacturing sector and using a CES demand system). Column (3) differs from the benchmark model insofar as firms decide in the current period, after current productivity is realized, whether to pay the fixed cost and whether (and how much) to automate, such that new capital technologies are immediately productive. Column (4) differs from the benchmark insofar as the task-substitution elasticity is lower (while, again, other parameters are re-calibrated to match the same data targets). Column (5) features a different parameterization of the capital and labor productivity schedules, which allows for a lower long-run capital-labor elasticity η_{LR} (the parameter γ refers to a different object and is not directly comparable to other model versions). η_{LR} is not a constant in this exercise. Locally, at the initial steady state, it equals 0.90, and it falls only slightly to 0.88 over the range considered in this exercise.

C.2 Short-run capital-labor elasticity

We argued in the main text that we view the calibrated short-run elasticity of $\eta = 0.4$ as an upper bound on the elasticity of substitution between tasks, since the firm-level evidence in Oberfield and Raval (2014) refers to cross-sectional variation, which might be interpreted as incorporating adjustments in task allocations within firms. Column (4) in Table 7 describes an alternative calibration with $\eta = 0.2$. For this exercise, we hold the long-run elasticity constant, which requires increasing the comparative advantage

TABLE 8: Model robustness: Transitional dynamics

	DATA	MODEL			
		BENCHMARK	ALTERN. TIMING	LOWER η	$\eta_{LR} < 1$
	(1)	(2)	(3)	(4)	(5)
<i>I. Parameters and inferred aggregate shocks</i>					
$d \ln q$.	1.40	1.40	1.39	3.80*
c_a	.	0.23	0.18	0.24	0.35
<i>II. Targeted moments, 1982–2012</i>					
Δ aggregate labor share	−0.178	−0.176	−0.181	−0.176	−0.174
Δ median labor share	0.030	0.026	0.031	0.029	0.016
<i>III. Other moments, 1982–2012</i>					
Δ Top 4 firms’ sales share	0.060	0.043	0.045	0.029	0.222
Δ Top 20 firms’ sales share	0.052	0.070	0.074	0.049	0.207
Δ log productivity dispersion	0.050	0.057	0.058	0.059	0.015
<i>IV. Melitz–Polanec decomposition from Autor et al. (2020)</i>					
Δ aggregate labor share	−0.185	−0.176	−0.181	−0.176	−0.174
Δ unweighted mean	−0.002	0.032	0.038	0.035	0.002
Exit	−0.055	−0.003	−0.003	−0.003	−0.010
Entry	0.059	0.006	0.005	0.006	0.016
Covariance term	−0.187	−0.212	−0.221	−0.215	−0.183
<i>V. Covariance decomposition from Kehrig and Vincent (2020)</i>					
Market share dynamics	0.047	0	0	0	0
Labor share by size dynamics	−0.043	−0.036	−0.027	−0.036	−0.041
Joint dynamics	−0.232	−0.143	−0.157	−0.144	−0.140

Notes: See Section 1 and Table 2 for details on the benchmark model (corresponding to the manufacturing sector and using a CES demand system). Column (3) differs from the benchmark model insofar as firms decide in the current period, after current productivity is realized, whether to pay the fixed cost and whether (and how much) to automate, such that new capital technologies are immediately productive. Column (4) differs from the benchmark insofar as the task-substitution elasticity is lower. Column (5) features a different parameterization of the capital and labor productivity schedules, which allows for a lower long-run capital-labor elasticity. The calibrated shock $d \ln q$ is not comparable to other model versions, as it does not raise investment efficiency uniformly across all tasks (see details in text).

parameter γ from $1.35 - 0.4 = 0.95$ to $1.35 - 0.2 = 1.15$. Comparing columns (2) and (4) in Table 8 reveals that the results virtually coincide, implying that our findings are not sensitive to the value of η within a reasonable range.

C.3 Long-run capital-labor elasticity

In the main text, we studied the effects of a uniform decline in the price of capital goods. This shock requires an above one long-run capital-labor elasticity ($\eta_{LR} > 1$) to match the decline in the labor share observed in the data. Following Proposition 3, this appendix shows that even if $\eta_{LR} < 1$, we can also generate a decline in the aggregate labor share and an increase in the labor share of the typical median firm by introducing declines in the price of capital at marginal tasks.

Implementing this exercise requires using a different parametrization of the schedules of capital and labor productivity, since the one used in the main text does not allow for $\eta_{LR} < 1$.³⁵ Specifically, we use the parametrization

$$\psi_k(x) = \begin{cases} (x_L - x)^\gamma & \text{if } x \leq x_L \\ 0 & \text{if } x > x_L \end{cases}$$

$$\psi_\ell(x) = x^\gamma,$$

where $x_L \in (0, 1]$ denotes a cut-off such that all tasks in $(x_L, 1]$ can only be performed by labor, and γ denotes the comparative advantage parameter. Setting $x_L = 0.9$ and $\gamma = 5$ generates a long-run elasticity that ranges between 0.88 and 0.90 along the transitional dynamics studied in this exercise. Column (5) in Table 7 describes the full steady state calibration.

With this parametrization, the aggregate labor share would increase in both the short and long run in response to a uniform increase in $q(x)$. Instead, we consider a task-specific increase in $q(x)$ of the following form:

1. From 1982, when the economy is in the initial steady state, until 1992, $\ln q(x)$ increases linearly for all $x \geq \alpha_{1982}^*$.
2. Over 1992-2002, $\ln q(x)$ increases linearly for all $x \geq \alpha_{1992}^*$.
3. Over 2002-2012, $\ln q(x)$ increases linearly for all $x \geq \alpha_{2002}^*$.

The particular formulation of the shock is not crucial; the main point is that we want to engineer an increase in capital productivity at marginal tasks while limiting a reduction of capital prices at infra-marginal tasks that are already performed by capital. This reduction reduces the capital share at those tasks, and is the reason why a uniform decline in the price of capital across all tasks requires $\eta_{LR} > 1$ to generate a decline in the aggregate labor share.

Column (5) in Table 8 describes the transitional dynamics in this case. For some tasks, we infer a change in log investment efficiency of up to 3.8, but one should keep in mind that this does not map to the observed decline in the price of any broad category of capital goods.³⁶ The inferred fixed cost of automation needs to be around 50% higher than in

³⁵The specification in the main text is the unique specification with the property that η_{LR} is a constant, but it requires $\eta_{LR} > 1$. All other specifications, and by implication all specifications that allow for $\eta_{LR} < 1$, are such that η_{LR} is not a constant.

³⁶For $x \in [0, \alpha_{1982}^*]$, $q(x)$ does not change at all; for $x \in [\alpha_{1982}^*, \alpha_{1992}^*]$, $\ln q(x)$ increases linearly by 1.27

the benchmark to generate a slight increase in median and unweighted mean labor shares, since the lack of an increase in capital productivity for tasks that are initially performed by capital implies that the labor shares of non-automating firms increases by little. While the labor share decompositions are comparable to the main exercise, productivity dispersion and sales concentration (Panel III) react differently: Since capital productivity increases primarily in marginal tasks, the productivity benefit of automation is smaller. Thus, first, the rise in productivity dispersion amounts to merely a quarter to one third of the effect generated in the benchmark exercise (which roughly matches the data). Second, the fraction of automating firms is much smaller, such that only the top firms automate at all. As a result, the difference between the top and the typical firm is much bigger, and the rise in sales concentration is much stronger. We conclude from this exercise that, even when the long-run elasticity of substitution is below 1, we can generate the observed decline in the aggregate labor share and the rise in the labor share of the typical firm as a response to a reduction of capital prices at marginal tasks.

D CALIBRATION OF THE NON-CES DEMAND MODEL FOR OTHER SECTORS

Table 9 summarizes the steady state calibration of the model with size-dependent markups in the wholesale as well as the utilities & transportation sector. The calibration strategy is identical to manufacturing and retail, which we describe in the main text. The log-convexity of the z -distribution is rather mild in these two sectors (n only slightly below 1), more in line with manufacturing than with retail.

Table 10 shows the model-based decomposition exercise, where we follow the same strategy as for manufacturing and retail. In wholesale as well as in utilities & transportation, the labor share decline is of the same magnitude as in retail, while the observed increase in concentration is not as drastic. Consequently, the inferred decline in the price of capital ($d \ln q$) is of similar magnitude, while the inferred increase in competition ($d \ln \lambda$) is weaker than in retail. The inferred automation fixed costs (c_a) are small, both in absolute terms and relative to manufacturing and retail. Comparing the various model versions, we find that the declining capital price caused 8.1pp or 70% of the sectoral labor share decline in wholesale, and 5.2pp or 72% of the sectoral labor share decline in utilities & transportation. For both sectors, the remainder is almost entirely due to the non-

from 1982 to 1992; for $x \in [\alpha_{1992}^*, \alpha_{2002}^*]$, $\ln q(x)$ increases linearly by 2.53 from 1982 to 2002; and for $x \in [\alpha_{2002}^*, 1]$, $\ln q(x)$ increases linearly by 3.8 from 1982 to 2012.

TABLE 9: Steady state calibration of the non-CES demand model: Wholesale, Utilities & Transportation

	PARAMETER		MOMENT	DATA	MODEL
<i>I. Wholesale: steady state parameters and moments (1982)</i>					
$\ln q_0$	Inverse capital price	-5.97	Aggregate labor share	58.6%	58.6%
σ	Demand elasticity	9.2	Aggregate markup	1.15	1.15
ν/σ	Demand supra-elasticity	0.20	Median labor share ratio	1.169	1.130
ζ	Weibull scale	0.098	Top 20 firms' sales share	42.9%	42.9%
n	Weibull shape	0.83	Top 4 firms' sales share	22.3%	22.3%
\underline{c}_f	Minimum fixed cost	$6.5 \cdot 10^{-8}$	Entry (=exit) rate	0.062	0.063
ξ	Dispersion fixed cost	0.210	Size of exiters	0.490	0.485
μ_e	Entrant productivity	0.888	Size of entrants	0.600	0.596
<i>II. Utilities & Transportation: steady state parameters and moments (1992)</i>					
$\ln q_0$	Inverse capital price	-5.38	Aggregate labor share	52.0%	52.0%
σ	Demand elasticity	10.2	Aggregate markup	1.15	1.15
ν/σ	Demand supra-elasticity	0.16	Median labor share ratio	1.169	1.133
ζ	Weibull scale	0.143	Top 20 firms' sales share	59.1%	56.1%
n	Weibull shape	0.95	Top 4 firms' sales share	30.4%	33.1%
\underline{c}_f	Minimum fixed cost	$7.0 \cdot 10^{-9}$	Entry (=exit) rate	0.062	0.062
ξ	Dispersion fixed cost	0.180	Size of exiters	0.490	0.488
μ_e	Entrant productivity	0.888	Size of entrants	0.600	0.600

Notes: The aggregate labor shares correspond to the BLS Multifactor Productivity Tables estimate for the wholesale as well as utilities & transportation sector. The ratio median-to-aggregate is from Kehrig and Vincent (2020), which refers to manufacturing, since in these sectors the census data used, e.g., by Autor et al. (2020) does not allow to compute the labor share of value added. The two concentration measures are from Autor et al. (2020) and correspond to these two sectors in 1982, respectively 1992. The model equivalents refer to the top 0.074% and top 0.369% of firms ranked by sales in wholesale (since there are on average 5,420 firms per 4-digit wholesale industry). For utilities & transportation, the model equivalents correspond to the top 0.100% and top 0.499% of firms ranked by sales (since there are on average 4,010 firms per 4-digit industry in this sector). The data moments and calibration strategy on entry rates, size of exiters and of entrants, follow the model with CES demand, see Table 1. In each of the two sectors, the eight parameters are jointly calibrated to match the eight data moments.

linearity or interaction effect of the two shocks: rising competition increases automation incentives for the top firms, and automation increases labor share differentials, magnifying the effect of reallocation on the aggregate labor share decline. We conclude that while for these two sectors, both shocks are important to account for the labor share decline, the inferred q -shock by itself is relatively more important than in retail, and less than in manufacturing.

TABLE 10: Transitional dynamics and decomposition of the labor share using a non-CES demand system: Wholesale, Utilities & Transportation

	DATA	MODEL		
		BENCHMARK	ONLY EFFECTS OF $d \ln q$	ONLY EFFECTS OF $d \ln \lambda$
	(1)	(2)	(3)	(4)
A. Wholesale (1982–2012)				
<i>I. Parameters and inferred aggregate shocks</i>				
$d \ln q$.	0.74	0.74	0
$d \ln \lambda$.	0.27	0	0.27
c_a	.	0.05	0.05	0.05
<i>II. Targeted moments, 1982–2012</i>				
Δ aggregate labor share	−0.114	−0.115	−0.081	−0.001
Δ unweighted mean labor share	0.047	0.050	0.024	0.007
Δ Top 4 firms' sales share	0.047	0.063	0.002	0.058
Δ Top 20 firms' sales share	0.101	0.086	0.003	0.080
<i>III. Other moments, 1982–2012</i>				
Δ log productivity dispersion	.	0.026	0.014	0.000
<i>IV. Markups, 1982–2012</i>				
Δ log aggregate markup	.	0.003	0.001	0.001
Within firm change in markups	.	−0.023	−0.010	−0.018
Reallocation	.	0.025	0.011	0.019
B. Utilities & Transportation (1992–2012)				
<i>I. Parameters and inferred aggregate shocks</i>				
$d \ln q$.	0.40	0.40	0
$d \ln \lambda$.	0.22	0	0.22
c_a	.	0.002	0.002	0.002
<i>II. Targeted moments, 1992–2012</i>				
Δ aggregate labor share	−0.073	−0.072	−0.052	0.000
Δ unweighted mean labor share	0.004	0.007	−0.002	0.003
Δ Top 4 firms' sales share	0.046	0.045	0.000	0.044
Δ Top 20 firms' sales share	0.047	0.049	−0.001	0.050
<i>III. Other moments, 1992–2012</i>				
Δ log productivity dispersion	.	0.015	0.007	0.000
<i>IV. Markups, 1992–2012</i>				
Δ log aggregate markup	.	0.000	0.000	0.000
Within firm change in markups	.	−0.011	−0.004	−0.010
Reallocation	.	0.011	0.004	0.010

Notes: Column (2) reports the findings from our benchmark model, which calibrates a uniform decline in the capital price, an increase in competition, as well as the automation fixed cost to replicate the change in the aggregate sectoral labor share (BLS), the unweighted mean labor share change among surviving firms (Autor et al., 2020, Table 5), and the increase in the top 4 as well as top 20 firms' sales share (Autor et al., 2020, Table 1), for both sectors. Due to data availability, the transition is over 1982–2012 for wholesale, resp. 1992–2012 for utilities & transportation. Column (3) shows results when shutting down the competition shock, and column (4) when shutting down instead the price of capital shock. Panel IV displays the log change in the aggregate markup, as well as a decomposition into within firm and reallocation components as in Baqaee and Farhi (2020b).

E COMPARING ESTIMATED SHOCKS TO DATA

This section provides additional motivation for our focus on the 1982–2012 period and benchmarks the inferred shocks and calibrated model parameters to the available data.

Historical behavior of payroll shares: As a starting point, Figure 10 provide data on payroll shares by sector for 1947–1987 and 1987–2012 from the BEA industry accounts. We split the data in these two periods due to changes in industry definitions introduced by the BEA in 1987, as it switched from the *Standard Industry Classification* to the *North American Industry Classification System*. We also focus on payroll shares, since labor shares of value added (which also include non-wage compensation), are not available for the earlier period. Figure 1 in the Introduction already provided the BLS data for labor shares of value added for the later period.

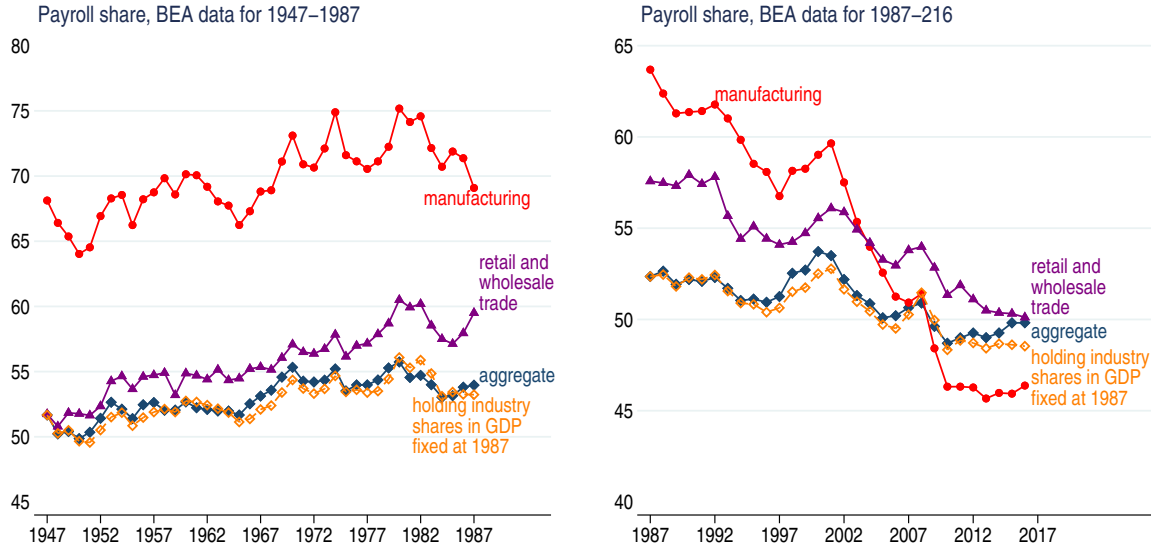


FIGURE 10: PAYROLL SHARE IN THE U.S. FOR 1947–2016. The figure plots the payroll share of value added for 1947–1987, both for some specific sectors and the economy as a whole. Data from the BEA industry accounts.

As discussed in the main text, the figure shows that payroll shares were constant or increasing up to 1982, and then started a sharp decline both in manufacturing, retail and wholesale. This motivates our focus on the 1982–2012 period and supports our choice of 1982 as the steady state of the model.

Benchmarking the fixed cost of automating tasks: As discussed in the main text, the fixed cost of automating tasks can be thought of as an investment in R&D required

to design and integrate automation equipment or software. As such, these fixed costs will contribute to rising R&D expenditures in the economy. The left panel in Figure 11 compares the inferred behavior of automation fixed costs to the available data on R&D spending. In particular, the panel displays the time series of automation cost spending as a share of output in the model, $\int c_a \cdot \max\{0, \hat{\alpha}_t(z) - \alpha\} \cdot m_t(\theta) d\theta$, both for manufacturing and for retail. For simplicity, we focus on the non-CES version of the model calibrated in Section 2.4. As the price of capital declines from 1982 to 2012 and stays constant thereafter, automation cost spending in manufacturing reaches a peak of roughly 0.7% of output around 2005, and declines to zero eventually as the model economy converges to the new steady state. In retail, automation cost spending is much lower, around 0.1% of output in 2005.

The black line provides the behavior of R&D as a share of GDP for the entire US economy. This share rose from a level of 1% before 1980 to a current level of 2.5% of GDP. Because not all R&D expenditure is due to automation, we see this series as an upper bound for automation cost spending. In line with this view, our calibration implies that about 25% of manufacturing R&D after 1980 corresponds to investments in automation fixed costs, and that rising automation cost spending since 1980 explain about 0.6 percentage points of the 1 percentage point increase in R&D observed from 1980–2005. This comparison shows that our estimated fixed costs are of a reasonable magnitude, and that they do not generate a counterfactual increase in R&D spending.

Benchmarking the inferred decline in capital prices and rising competition:

The right panel of Figure 11 compares the inferred decline in the price of capital $d \ln q$ to data. As the empirical counterpart for these series, we use the percent decline in the price of equipment and software over the time period 1982–2012 from the BEA’s Fixed Asset Tables, which we deflate using the PCE index. We focus on software and equipment because these are most relevant for capital–labor substitution in our framework. In addition, we display a series by DiCecio (2009), which attempts to correct the BEA series for missing quality-adjustment in the spirit of Gordon (1990), following the imputation procedure in Cummins and Violante (2002).

The inferred decline in the price of capital in our model is well within the range of the empirical counterparts. In particular, the decline in the price of equipment and software needed to explain the decline of the manufacturing labor share is of 140 log points, which is comparable to the observed decline of 174 log points in the data (112 without Gordon’s quality adjustment).

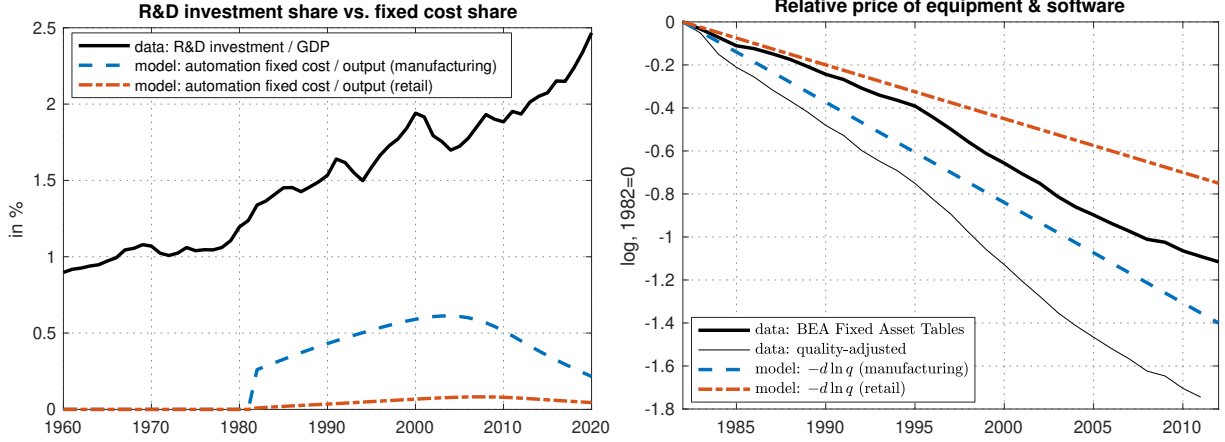


FIGURE 11: AUTOMATION FIXED COSTS AND CAPITAL PRICE DECLINES. The left panel displays the time series of automation fixed cost spending relative to aggregate output for manufacturing and retail using our estimates from Section 2.4. Before 1982, the model is in steady state, and automation spending is zero. The black line corresponds to the ratio of R&D investment spending relative to U.S. GDP. The right panel displays the calibrated decline in capital prices $-d \ln q$ for manufacturing and retail using our estimates from Section 2.4. The black line plots the observed decline in the price of equipment and software capital for 1982–2012, deflated by the PCE price index (source: BEA Fixed Asset Tables). In addition, the gray line plots another series for the relative price of equipment and software due to DiCecio (2009), which builds on work by Gordon (1990) and Cummins and Violante (2002) and imputes missing quality-adjustment.

Turning to the inferred measure of rising competition, it is difficult to find an empirical counterpart to λ in the data. Our model infers a particular strong increase in retail ($d \ln \lambda = 0.41$), followed by wholesale (0.27), utilities/transportation (0.22), and manufacturing (0.04). The type of changes that λ proxies for include in particular a rise in the effective market size caused by the widespread availability of internet search engines, and in general breakthroughs in information and communications technology, which reduce information frictions. For example, Akerman, Leuven and Mogstad (2021) provide evidence showing that broadband availability is causally associated with an expansion of the choice set of importers and exporters in Norway. At a qualitative level, it is reasonable that this type of technological change most strongly affected retail trade, followed by wholesale trade.

F COMPUSTAT DATA AND ADDITIONAL EMPIRICAL RESULTS REGARDING MARKUPS AND OUTPUT ELASTICITIES

F.1 Data description, sample, and definitions

We use data from Compustat from 1960 to 2016. We use the following variable definitions and conventions:

- Revenue y_{ft} : we measure revenue using firm sales—*SALES* in Compustat.
- Expenditures in variable inputs v_{ft} : we measure these expenditures using the cost of goods sold—*COGS* in Compustat.
- Stock of capital k_{ft} : we measure capital using the gross value of property, plants, and equipment—*PPEGT* in Compustat.
- Investment x_{ft} : we measure investment as the percent change in capital; that is, $\ln x_{ft} = \ln k_{ft+1} - \ln k_{ft}$
- Industry and firm groupings $c(f)$: we conduct our estimation separately for 23 NAICS industries, roughly defined at the 2-digit level. When grouping firms into size quintiles, we do so for each year and within each 3-digit NAICS industry. We also experimented with the classification of industries based on SIC codes used in Baqaee and Farhi (2020b) and obtained very similar results.
- Sample definition and trimming: following De Loecker, Eeckhout and Unger (2020), we trim the sample by removing firms in the bottom 5th and top 5th percentiles of the *COGS*-to-*SALES* distribution. In addition, following Baqaee and Farhi (2020b), we exclude firms in farm and agriculture, construction, real estate, finance, and utilities from our markup and labor share calculations in Figures 7 to 9.
- Winsorizing: we winsorize the obtained revenue elasticities at zero. Moreover, following Baqaee and Farhi (2020b), we winsorize our markup estimates at the 5th and 95th percentile of their distribution.

F.2 Robustness checks

This sub-section provides estimates for firm markups using different sets of assumptions.

Estimates parametrizing g and h using cubic polynomials We estimate elasticities under the same assumptions outlined in the main text, but parametrize g and h using cubic polynomials. Figure 12 plots the behavior of the resulting output elasticities over time by firm size quintile. Figure 13 reports the contribution of within-firm changes in markups and between-firm reallocation to (percent) changes in the labor share.

Estimates assuming there are no ex-post shocks ϵ In the absence of ex-post shocks, we can treat observed revenue as true revenue and there is no need to use a proxy variable to recover productivity. Instead, we can compute productivity directly as

$$z_{ft} = \ln y_{ft} - \varepsilon_{vc(f)t}^R \cdot \ln v_{ft} - \varepsilon_{kc(f)t}^R \cdot \ln k_{ft},$$

and proceed with the rest of the estimation in the same way as before.

Figure 14 plots the behavior of the resulting output elasticities over time by firm size quintile. Figure 15 reports the contribution of within-firm changes in markups and between-firm reallocation to (percent) changes in the labor share.

Estimates assuming a linear Markov process for productivity Suppose that productivity follows a linear Markov process

$$z_{ft} = \beta z_{ft-1} + \zeta_{ft}.$$

Define $v_{ft} := z_{ft} + \epsilon_{ft}$. Because ex-post shocks are i.i.d, we have that v_{ft} also follows a linear Markov process

$$v_{ft} = \beta v_{ft-1} + \underbrace{\zeta_{ft} + \epsilon_{ft} - \beta \epsilon_{ft-1}}_{:=\iota_{ft}}.$$

Estimation proceeds as follows. First, we can compute v_{ft} directly as

$$\tilde{v}_{ft} = \ln y_{ft} - \varepsilon_{vc(f)t}^R \cdot \ln v_{ft} - \varepsilon_{kc(f)t}^R \cdot \ln k_{ft}.$$

Then we estimate the linear model

$$\tilde{v}_{ft} = \beta \tilde{v}_{ft-1} + \tilde{\iota}_{ft},$$

and base estimation on the moment conditions

$$\mathbb{E}[\iota_{ft} \otimes (\ln k_{ft}, \ln v_{ft-1})] = \mathbb{E}[(\zeta_{ft} + \epsilon_{ft} - \beta \epsilon_{ft-1}) \otimes (\ln k_{ft}, \ln v_{ft-1})] = 0.$$

Figure 16 plots the behavior of the resulting output elasticities over time by firm size quintile. Figure 17 reports the contribution of within-firm changes in markups and between-firm reallocation to (percent) changes in the labor share.

F.3 Biases of a sales weighted average of markups

This subsection demonstrates that, when there are differences in output elasticities across firms that are not taken into account, a sales-weighted average of firm-level markups provides a biased assessment of the contribution of changes in markups to the decline in the labor share. First, as demonstrated in the main text (see footnote 27), the true contribution of markups to the labor share is given by

$$\mu_t := 1 / \sum_f \omega_{ft} \cdot \frac{1}{\mu_{ft}}.$$

Suppose that, as is typically done in the literature, we compute firm-level markups under the assumption that all firms operate a technology with common output elasticities, ε_{vt} . This implies that the measured firm-level markups μ_{ft}^m and the true markups μ_{ft} satisfy

$$\mu_{ft}^m = \frac{\varepsilon_{vt}}{\varepsilon_{vft}} \cdot \mu_{ft},$$

where ε_{vft} is the firm-level elasticity of output with respect to variable input. This implies that measured firm-level markups will over-state the true markups when $\varepsilon_{vt} > \varepsilon_{vft}$ and under-state them otherwise. Moreover, the average elasticity is given by

$$\varepsilon_{vt} = \sum_f \omega_{ft} \cdot \varepsilon_{vft}.$$

We now show that these firm-level biases do not cancel when one computes an arithmetic sales-weighted average of markups, and that this problem aggravates when there dispersion of ε_{vft} rises.

The measured sales-weighted markup is given by

$$\bar{\mu}_t^m = \sum_f \omega_{ft} \cdot \mu_{ft}^m.$$

After some algebra, one can show that

$$\frac{\bar{\mu}_t^m}{\mu_t} = 1 + \frac{\sum_{f,f'} y_{ft} \cdot y_{f't} \cdot \left(\sqrt{\frac{s_{vf't}}{s_{vft}}} - \sqrt{\frac{s_{vft}}{s_{vf't}}} \right)^2}{\left(\sum_f y_{ft} \right)^2}.$$

This implies that the measured sales weighted markup exceeds the true markup, with equality if and only if there is no dispersion in factor shares $s_{vft} = s_{vf't}$.

Moreover, this expression shows that the gap between the measured sales-weighted markup and the real markup rises with the dispersion in factor shares across firms, even if this dispersion is fully driven by differences in technology. In particular, if all firms had $\mu_{ft} = 1$ and charged no markups, a sales-weighted average of markups computed under the assumption of common technology would be equal to

$$\bar{\mu}_t^m = 1 + \frac{\sum_{f,f'} y_{ft} \cdot y_{f't} \cdot \left(\sqrt{\frac{\varepsilon_{vf't}}{\varepsilon_{vft}}} - \sqrt{\frac{\varepsilon_{vft}}{\varepsilon_{vf't}}} \right)^2}{\left(\sum_f y_{ft} \right)^2}.$$

This sales-weighted markup not only exceeds 1, but would also increase spuriously with the dispersion in output elasticities, ε_{vft} . Thus, the uneven adoption of capital intensive technologies, like the one produced by our model along the transition, can generate a spurious increase in the sales-weighted markup, even if there are no changes in market power.

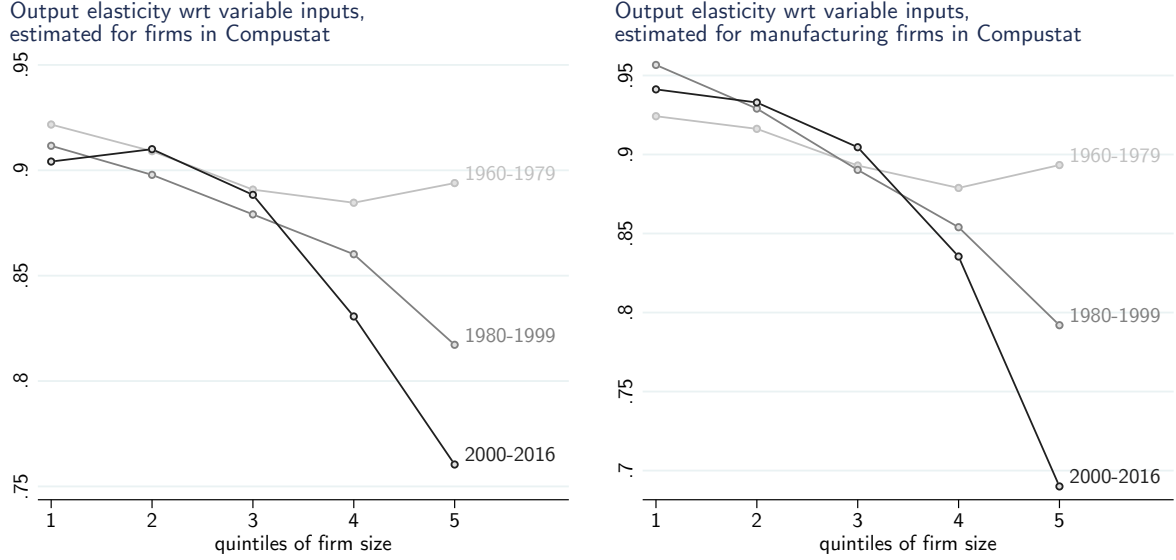


FIGURE 12: OUTPUT ELASTICITY WITH RESPECT TO VARIABLE INPUTS ESTIMATED USING A CUBIC PARAMETRIZATION OF g AND h . The left panel presents estimates obtained for all Compustat firms. The right panel presents estimates obtained for Compustat firms in manufacturing. See figure 7 in the main text for our baseline estimates used in the results reported in the paper.

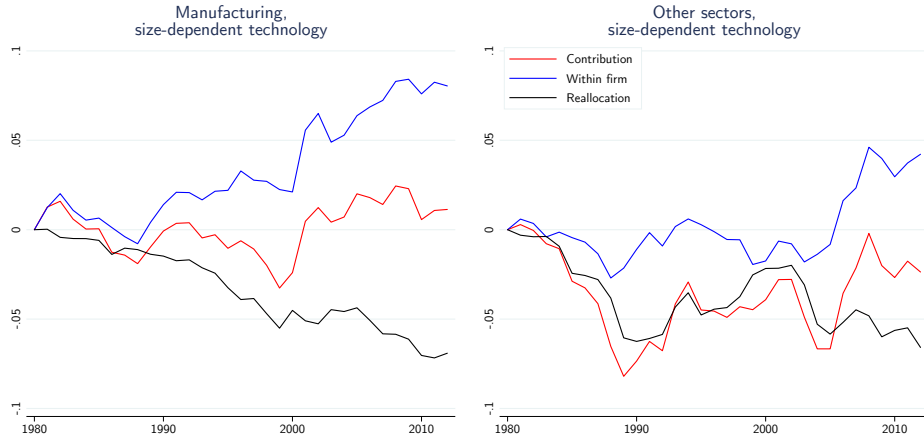


FIGURE 13: DECOMPOSITION OF THE CONTRIBUTION OF WITHIN-FIRM CHANGES IN MARKUPS AND BETWEEN-FIRM REALLOCATION TO (PERCENT) CHANGES IN THE LABOR SHARE. See the main text for details on this decomposition. Firm-level markups are estimated using a cubic parametrization for g and h , as explained in Appendix F.2. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other economic sectors. See figure 9 in the main text for our baseline estimates reported in the paper.

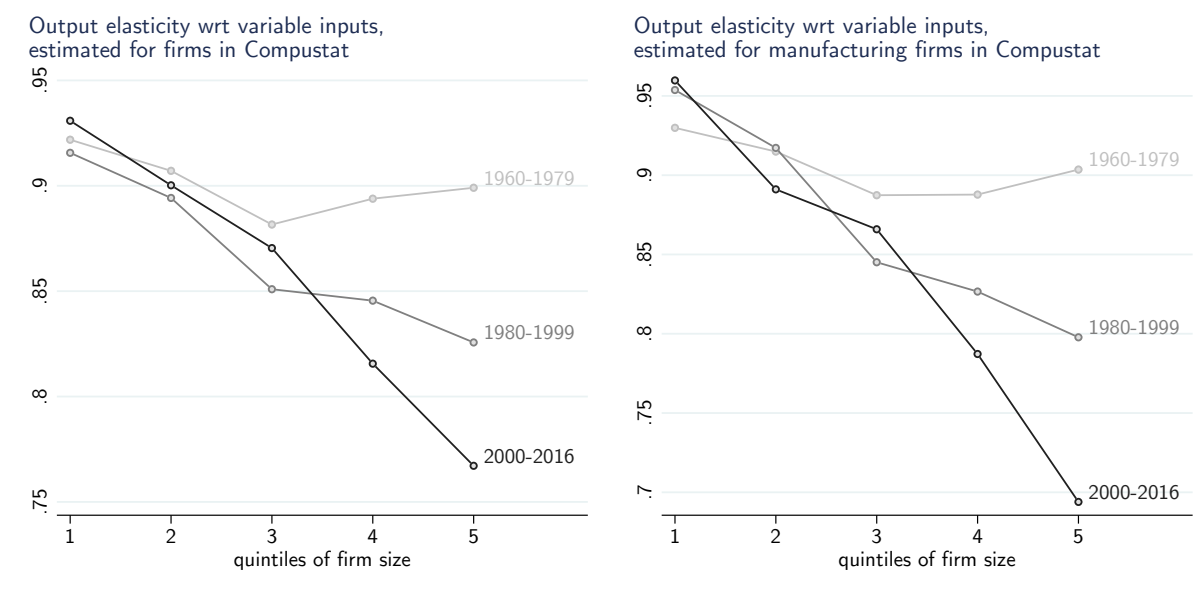


FIGURE 14: OUTPUT ELASTICITY WITH RESPECT TO VARIABLE INPUTS ESTIMATED UNDER THE ASSUMPTION THAT THERE ARE NO EX-POST SHOCKS. The left panel presents estimates obtained for all Compustat firms. The right panel presents estimates obtained for Compustat firms in manufacturing. See figure 7 in the main text for our baseline estimates used in the results reported in the paper.

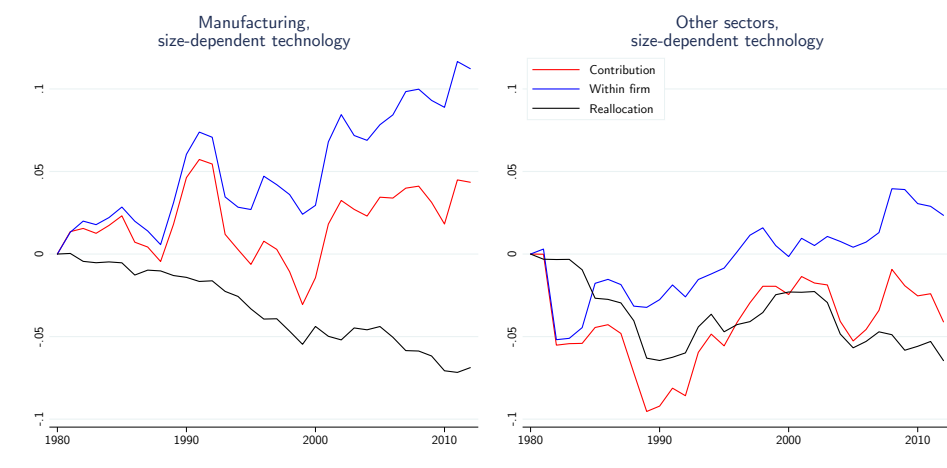


FIGURE 15: DECOMPOSITION OF THE CONTRIBUTION OF WITHIN-FIRM CHANGES IN MARKUPS AND BETWEEN-FIRM REALLOCATION TO (PERCENT) CHANGES IN THE LABOR SHARE. See the main text for details on this decomposition. Firm-level markups are estimated under the assumption of no ex-post shocks, as explained in Appendix F.2. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other economic sectors. See figure 9 in the main text for our baseline estimates reported in the paper.

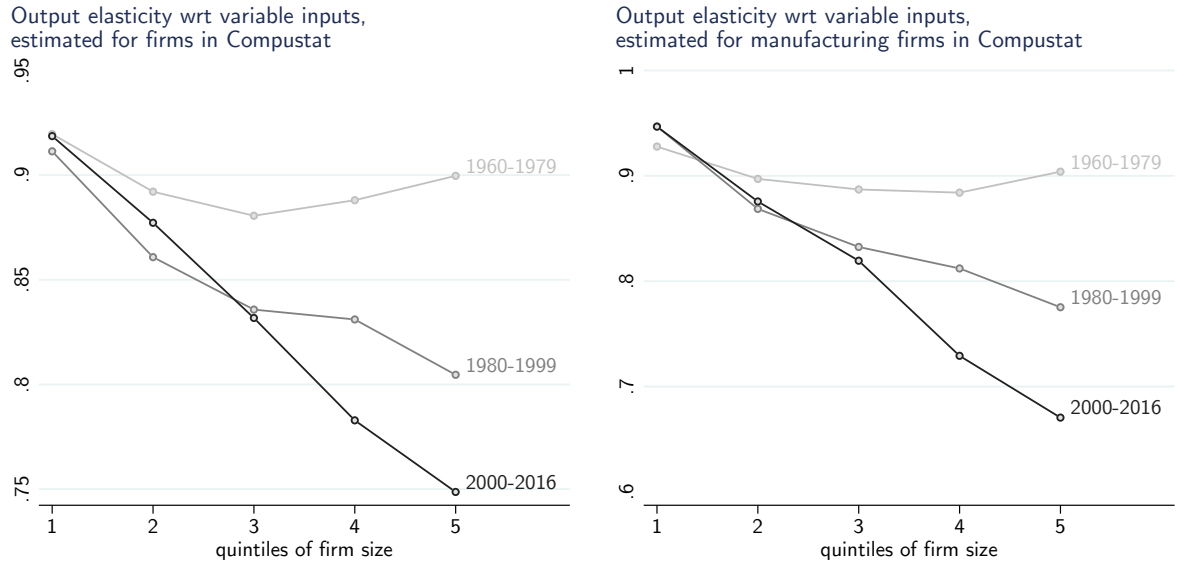


FIGURE 16: OUTPUT ELASTICITY WITH RESPECT TO VARIABLE INPUTS ESTIMATED UNDER THE ASSUMPTION THAT PRODUCTIVITY FOLLOWS A LINEAR MARKOV PROCESS. The left panel presents estimates obtained for all Compustat firms. The right panel presents estimates obtained for Compustat firms in manufacturing. See figure 7 in the main text for our baseline estimates used in the results reported in the paper.

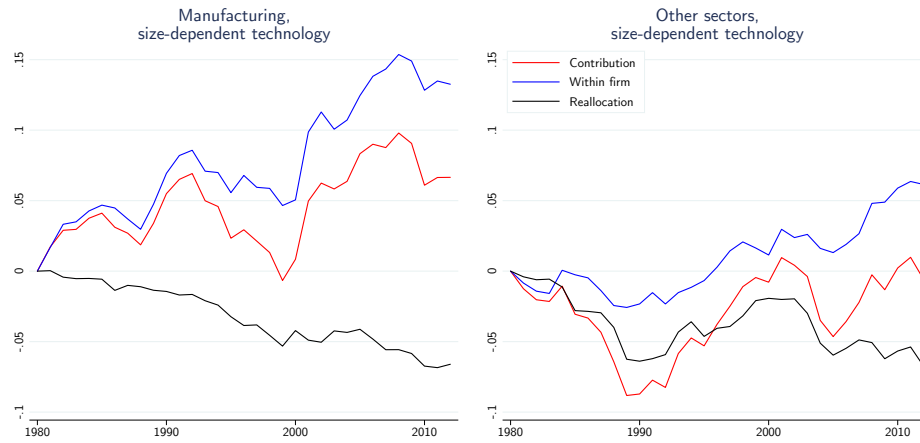


FIGURE 17: DECOMPOSITION OF THE CONTRIBUTION OF WITHIN-FIRM CHANGES IN MARKUPS AND BETWEEN-FIRM REALLOCATION TO (PERCENT) CHANGES IN THE LABOR SHARE. See the main text for details on this decomposition. Firm-level markups are estimated under the assumption that productivity follows a linear Markov process, as explained in Appendix F.2. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other economic sectors. See figure 9 in the main text for our baseline estimates reported in the paper.