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, Revised January 2022

We thank Daron Acemoglu, David Autor, Hugo Hopenhayn, Loukas Karabarbounis, Dirk Krueger, Stephen Terry, as well as conference and seminar participants at Bank of Mexico, Barcelona GSE Summer Forum, Boston College, Columbia University Junior Conference, Copenhagen, Fed Richmond, IMF, MMF Workshop on Macroeconomic Consequences of Technical Change, NBER Summer Institute, Oslo Macro Conference, Peking University, SED Meetings Minneapolis, Toronto, and Virtual Macro Seminar Series for their comments. Restrepo thanks the National Science Foundation for its support under award No. 2049427. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

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, Revised January 2022
JEL No. E22,E23,E24,E25

ABSTRACT

While the US labor share has declined, especially in manufacturing and retail, the labor share of a typical firm has increased. This paper introduces a model where firms incur fixed costs to automate tasks. In response to lower capital prices, the model reproduces the labor share dynamics observed in the data: large firms automate more tasks, reducing the aggregate labor share; while the median firm continues to operate a labor-intensive technology with a rising labor share. Using our model, we decompose the labor share decline and the rise in concentration into a part driven by lower capital prices and a part driven by sales reallocation to more productive higher-markup firms. Reallocation played a minor role in explaining the labor share decline in manufacturing but an important role in retail and other sectors during 1982–2012. These conclusions are in line with estimates of markups and capital elasticities from Compustat.

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 for much of the last century, the US labor share declined from a peak of 62% in the 1980s to 55% in 2012, as seen in Figure 1.[2] The decline of the labor share is not driven by changes in the industry composition of the US economy; if anything, holding industry shares in GDP constant at their 1982 levels leads to a more pronounced decline.[3] In fact, most of the fall in the US 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.

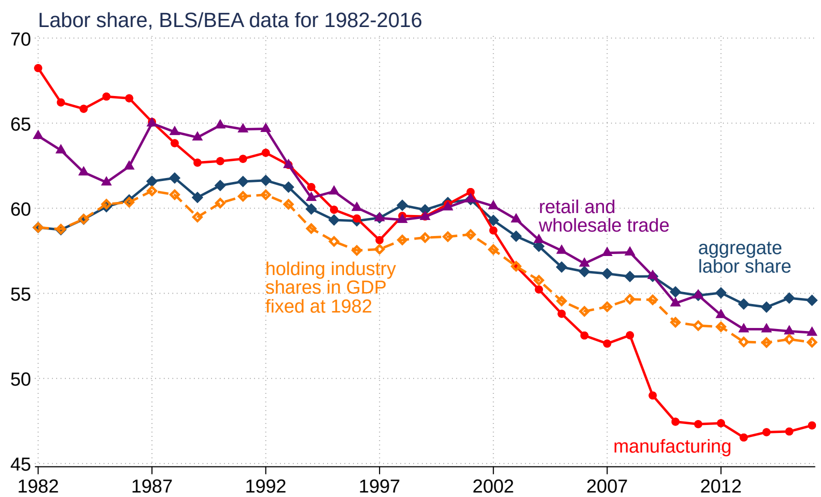

FIGURE 1: LABOR SHARE IN THE US The blue line plots the US labor share, excluding government and farming. The yellow line plots a counterfactual labor share holding industry shares in GDP constant at 1982 values. The red and purple lines plot the labor shares for manufacturing, and retail and wholesale trade. Data from the BEA-BLS integrated industry-level production account (Eldridge et al., 2020).

Despite growing consensus on these facts, there is ongoing debate about the causes of the decline. One set of explanations points to the increased substitution of capital for labor in the production of goods and services. In these accounts, the development of new or more efficient capital-intensive technologies 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 as a widening range of tasks are

[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 US phenomenon, see Gutierrez and Piton (2020).

[2]Although there is some consensus about the decline in the US labor share, there is debate on how the exact magnitude of the decline is affected by the treatment of self-employment (Elsby, Hobijn and Şahin, 2013), income shifting by business owners (Smith et al., 2019), and investment in software and intangible capital (Koh, Santaeulalia-Llopis and Zheng, 2020).

[3]See also Elsby, Hobijn and Şahin (2013); Acemoglu and Restrepo (2019); Hubmer (2020).

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 supports these explanations (Acemoglu and Restrepo, 2020; Acemoglu, Lelarge and Restrepo, 2020; Hubmer, 2020).

This evidence notwithstanding, recent studies using firm-level data show that the decline in labor shares is not uniform across firms. While the aggregate labor share has declined, the labor share of the *typical* US firm has increased or remained unchanged. In manufacturing, the median labor share across firms rose from 71% to 74% and the unweighted mean labor share decreased slightly (Autor et al., 2020; Kehrig and Vincent, 2020). A similar pattern holds for other sectors and is evident for other countries (Autor et al., 2020). At a first glance, these facts cast doubt on explanations based on capital–labor substitution, since a simple version of these theories in which firms face the same prices and have access to the same technologies implies a uniform decline in labor shares. Instead, these new facts favor a second set of explanations that emphasize the role of rising concentration and the reallocation of sales 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 industry have higher markups.[4]

This paper starts from the observation that the adoption of modern automation technologies is uneven and concentrates among large firms. For instance, using a new module in the *Annual Business Survey* covering 300,000 firms across all US economic sectors, Acemoglu et al. (2021) show that, even within detailed industries, firms in the top percentile of the employment size distribution are 1.7 times more likely to use industrial robots, specialized equipment and software, and artificial intelligence than firms between the 50th and 75th percentile. Dinlersoz and Wolf (2018) document a similar phenomenon for manufacturing firms using the *Survey of Manufacturing Technologies* from 1993.[5] This evidence suggests

[4]A different narrative is that, in some US industries, we have seen an erosion of competition due to weaker anti-trust enforcement, which could also lead to higher concentration and firms raising their markups (Philippon, 2019). However, the available estimates (as well as our own estimates in Section 3) suggest that the increase in markups is not due to a within-firm increase, but due to the reallocation across firms with different markups, which is in line with theories emphasizing the role of rising competition or winner-takes-all dynamics as the main driving force affecting markups.

A related narrative is that the decline in the labor share might reflect rising monopsony power. However, the available evidence suggests that labor market concentration has decreased over time, if anything raising the labor share by 3 percentage points since 1982 (Rossi-Hansberg, Sarte and Trachter, 2021; Berger, Herkenhoff and Mongey, 2019).

[5]This pattern is not unique to the US. Recent papers document that the adoption of industrial robots concentrates in the largest manufacturing firms across several countries (see Koch, Manuylov and Smolka, 2019; Humlum, 2019; Bonfiglioli et al., 2020; Acemoglu, Lelarge and Restrepo, 2020).

that large firms not only differ in their productivity (and potentially their markups) but also in the extent to which their production relies on capital-intensive technologies.

Our main point is that, once we account for the observed heterogeneity in technology, explanations of the labor share decline based on capital–labor substitution are consistent, both qualitatively and quantitatively, with the firm-level evidence in Autor et al. (2020) and Kehrig and Vincent (2020). Moreover, accounting for differences in technology among firms with different size modifies existing conclusions regarding the contributions of markups and technology to the decline in the labor share across sectors.

We develop this point in three exercises.

Section 1: Our first exercise shows that theories of capital–labor substitution can generate the labor share dynamics observed on the aggregate and across firms qualitatively and quantitatively. To do so, we build a standard firm-dynamics model with a CES demand structure augmented with automation decisions across tasks (as in Zeira, 1998; Acemoglu and Restrepo, 2018). The key assumption is that firms must make costly upfront investments to automate additional tasks.

We show analytically that, in response to lower capital prices, one could have a decline in the aggregate labor share at the same time as the median firm exhibits an increase in its labor share, and provide sufficient conditions for this outcome. The new mechanism driving this result is as follows: in response to a persistent decline in the price of capital, 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, due to the fixed cost of adoption, the median firm will not automate tasks and will continue to operate a labor-intensive technology. For the median firm, capital and labor remain complements, explaining why the labor share rises for the typical firm.[6]

We show that a version of this model where the fixed cost is calibrated to match the adoption rate gradient by size in Acemoglu et al. (2021) generates aggregate and firm-level labor share dynamics that match the manufacturing data. In particular, as capital prices decrease by 168 log points, both the median labor share and the unweighted mean labor share remain roughly unchanged, while the aggregate manufacturing labor share declines by 20 percentage points. 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

[6] A different mechanism going back to Houthakker (1955) emphasizes the possibility that a decline in capital prices reallocates economic activity towards more capital-intensive firms, even if firms do not change their factor intensities. Oberfield and Raval (2014) show that this reallocation has a small and negative effect on the aggregate labor share. Kaymak and Schott (2018) argue that this mechanism can explain a third of the decline in the manufacturing labor share in response to lower corporate taxes.

market share of firms and their labor share, and not by the change in the unweighted average of labor shares among incumbents. In turn, the negative covariance term reflects the fact that firms automate during periods of expanding sales, which aligns with the joint dynamics of changes in value added and labor shares for manufacturing firms documented in Kehrig and Vincent (2020). The uneven use of automation technologies also generates an endogenous rise in sales concentration which accounts for half of the observed change in manufacturing.

Section 2: Our second exercise extends our baseline model to quantify the contribution of two driving forces behind the decline of the labor share and the increase in sales concentration: i. a decline in capital prices leading to capital-labor substitution among large and growing firms—the technology view—, and ii. an increase in market size leading to rising competition and reallocation of sales towards more productive firms—the reallocation view.[7] For this exercise we also look at retail and other non-manufacturing sectors, where rising competition and winner-takes-all dynamics might be more prevalent.

We work with a non-CES demand system where markups increase with firm size and larger firms have lower passthroughs—Marshall's second law of demand.[8] In this model, both driving forces can generate rising sales concentration, a decline in the aggregate labor share, and a rise in the typical firm labor share. Separating their contribution across sectors becomes a quantitative question.

We show that, in each sector, one can recover the decline in capital prices and the increase in market size by exactly matching the observed change in sales concentration and the aggregate decline in the labor share. Two key observations explain this result. On the one hand, for our calibrated parameters across all sectors, an increase in market size increases sales concentration much more than it reduces the aggregate labor share. In particular, when the productivity distribution across firms is more log-convex than but not too far from Pareto, the decline in the labor share due to reallocation to high markup firms

[7] Previous works have studied other driving forces that could affect concentration and markups. Lashkari, Bauer and Boussard (2019); Aghion et al. (2019); De Ridder (2020); Mariscal (2020) study the role of ICT in allowing more productive firms to expand to new markets, achieve greater returns to scale, or reorganize their production hierarchies. Akcigit and Ates (2019) study the slowdown of technology diffusion from large to small firms, and Olmstead-Rumsey (2019) explores the implications of a decline in R&D productivity among small firms. Hopenhayn, Neira and Singhania (2018) explores the shifting age composition of US firms, and how this results in a reallocation to older firms with higher sales and lower labor shares. These driving forces differ from our mechanism, which applies more forcefully to manufacturing and emphasizes the fact that new automation technologies make large firms more capital-intensive.

[8] Non-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 systems and Marshall's laws.

and the decline in within firm markups generated by rising competition mostly offset each other and lead to a large increase in concentration accompanied by a small reduction in the sectoral labor share. On the other hand, for our calibrated fixed cost of task automation, a decline in capital prices reduces the sectoral labor share much more than it increases sales concentration. Because both shocks affect concentration and the labor share differentially, they can be recovered to exactly match trends in these two variables.

In retail (and similarly in other non-manufacturing sectors), rising competition accounts for up to 60% of the decline in the labor share between 1982–2012 and 90% of the increase in sales concentration. However, when looking at manufacturing this pattern reverses and capital–labor substitution explains over 90% of the decline in the labor share in this period and over 50% of the increase in concentration. The reason why rising competition plays a minor role in manufacturing is that, in this sector, concentration increased mildly relative to its sizable labor share decline (as evident from the data in Autor et al., 2020). In addition, based on the observed distribution of firm sales in this sector, we estimate a distribution of firm productivity that is close to Pareto. As a result, a large decline in capital prices is required to explain the patterns in manufacturing. Instead, the distribution of productivity for retail firms is more log-convex than in manufacturing. This observation combined with the fact that concentration rose sharply in retail leads to the conclusion that reallocation forces played a more prominent role in driving the labor share decline in this sector.

Section 3: Our third exercise looks at empirical estimates of output elasticities and markups among firms in Compustat. To capture differences in the use of automation technologies by firm size, we estimate flexible production functions where output elasticities are allowed to vary over time, by firm size class, and by industry. In line with the key mechanism in our model, we find that over time, the largest firms in each industry (especially in manufacturing) have experienced a large increase in their output-to-capital elasticity, indicating that their production processes have become more capital intensive. Turning to markups, we also reach a conclusion that supports our model results. Outside of manufacturing, reallocation of sales to high-markup firms explains half of the observed labor share decline. Instead, markups explain a small fraction of the labor share decline in manufacturing. Our results also show that, once we account for technology differences across size classes, the aggregate markup has been stable over time at 1.15–1.2.

1 Capital–labor substitution with adoption costs

We augment a firm-dynamics model (as in Hopenhayn, 1992) to include firms' decisions to automate tasks (as in Acemoglu and Restrepo, 2018). Our key innovation is to incorporate heterogeneity in the extent to which firms automate their production process, and to endogenize the evolution of these decisions as determined by the payment of a fixed cost per task. This fixed cost ensures that firms automate more tasks as they grow in scale, which allow us to match the evidence in Acemoglu et al. (2021).[9]

1.1 Model and theoretical properties

Environment: We consider an economy in discrete time indicated by the subscript t. Existing firms, f, produce differentiated varieties y_{tf} combined via a CES aggregator to produce a final good y_t, whose price we normalize to 1:

$$y_t = \left(\int_f y_{tf}^{\frac{\sigma-1}{\sigma}} \cdot df \right)^{\frac{\sigma}{\sigma-1}}.$$

Here $\sigma > 1$ denotes the elasticity of substitution across varieties. Firms are atomistic and charge a common and constant markup $\mu = \sigma/(\sigma - 1) > 1$.

Firms differ in their productivity z_{tf} and in the fraction of tasks or production processes they have automated, $\alpha_{tf} \in [0, 1]$. A firm produces output y_{tf} by combining a continuum of tasks indexed by x with task substitution elasticity $\eta \geq 0$:

$$y_{tf} = z_{tf} \cdot \left(\int_0^1 y_{tf}(x)^{\frac{\eta-1}{\eta}} \cdot dx \right)^{\frac{\eta}{\eta-1}}.$$

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

$$y_{tf}(x) = \begin{cases} \psi^k(x) \cdot k_{tf}(x) + \psi^\ell(x) \cdot \ell_{tf}(x) & \text{if } x \in [0, \alpha_{tf}] \\ \psi^\ell(x) \cdot \ell_{tf}(x) & \text{if } x \in (\alpha_{tf}, 1] \end{cases}$$

Here, $y_{tf}(x)$ denotes the quantity of task x, and $k_{tf}(x)$ and $\ell_{tf}(x)$ denote capital and labor

[9] The assumption that automating a task entails fixed costs is plausible and intuitive. Consider a car-manufacturing firm that wishes to automate welding. Besides purchasing the industrial robots required to complete this task, the firm must also hire a team of engineers and integrators to reorganize its plant and production processes, 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 in integration 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 rearranging its operations. These fixed costs represent an intangible investment (see Corrado, Hulten and Sichel, 2009, on the importance of intangibles), which increases firms' ability to use capital for additional tasks—an intangible asset.

employed to produce task x. Without loss of generality, we assume that

$$\frac{\psi^{\ell}(x)}{q_t(x)\cdot\psi^k(x)} \text{ is increasing in } x,$$

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

Firms face a competitive market for inputs. There is a fixed supply of labor ℓ rented to firms at a wage rate w_t. On the other hand, the capital used for task x is produced from the final good, with a unit of the final good yielding $q_t(x)$ units of capital. Capital is produced immediately each period and fully depreciates after use, which implies that its rental rate equals $1/q_t(x)$.

Incumbent firms begin a period with productivity z_{tf} and automation level α_{tf}. They then make optimal employment and capital utilization decisions and collect profits π_{tf}. Subsequently, firms draw a fixed operating cost $c_o \cdot y_t$, where $c_o \sim G(c_o)$, and decide whether to continue operating.[10] If they continue, they draw next period's productivity level $z_{t+1,f}$, which follows an exogenous first-order Markov process with $z_{t+1,f}$ increasing in z_{tf} (in a stochastic sense). We also assume that, for any increasing and unbounded function f, $\mathbb{E}[f(z_{t+1,f})|z_{tf}]$ converges to infinity when $z_{tf}\to\infty$ and to $f(0)$ when $z_{tf}\to 0$.

The key ingredient of our model are the endogenous automation decisions by firms. Incumbents can expand the set of automated tasks to include $(\alpha_{tf},\alpha_{t+1,f}]$ at a cost $c_a\cdot y_t\cdot(\alpha_{t+1,f}-\alpha_{tf})$, which implies a fixed cost of automation per task of $c_a\cdot y_t$. We also allow these technologies to diffuse gradually through the entry of new firms, as we explain next.

Every period a unit mass of potential entrants decides whether to enter the market. Entrants draw a productivity signal z from a distribution $G_e(z)$ and start with a common level of automation $\bar{\alpha}_t$. After observing z and the realization of the fixed operating cost c_o, entrants decide whether to pay the fixed cost and enter. We let the entry level of automation $\bar{\alpha}_t$ equal the unweighted average of α_{tf} among incumbents. This is a common specification used in models of technology diffusion (see Perla, Tonetti and Waugh, 2021), which offers a simple way to get automation technologies to diffuse over time, reflecting the standardization of these production techniques and the associated organizational changes required to deploy them. This diffusion simplifies our analytical characterization of the steady state, but is not required for our quantitative work.

Finally, when making entry and adoption decisions, incumbents and new entrants dis-

[10] 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. The assumption that the fixed cost is in units of the final good ensures that there is no mechanical relationship between firm size and its labor share. All differences in labor shares are therefore due to markups, technology or factor prices.

count the future at a constant interest rate r, which we take as exogenous. Throughout, we assume that $r > g_t$, where g_t is the growth rate of output between two consecutive periods.

Equilibrium: Denote by $p_{tf}(w)$ the price charged by a firm facing a wage w, by $c_{tf}(w)$ its cost, and by $\pi_{tf}(w)$ its profits. Given a path for investment productivities $q_t(x)$ and an initial distribution of firms $\{\alpha_{0f}, z_{0f}\}$, an equilibrium is given by a path for wages w_t and output y_t, and a path for the distribution of firms $\{\alpha_{tf}, z_{tf}\}$, such that for all $t \geq 0$:

E1. The ideal-price index condition holds

$$\int_f p_{tf}(w_t)^{1-\sigma} \cdot df = 1.$$

E2. The labor market clears

$$\int_f y_t \cdot p_{tf}(w_t)^{-\sigma} \cdot \frac{\partial c_{tf}(w_t)}{\partial w_t} \cdot df = \ell.$$

E3. Automation and exit decisions maximize the value function of incumbents[11]

$$\begin{aligned} V_{tf} =& \pi_{tf}(w_t) + \\ & \int \max \left\{ 0, -c_o \cdot y_t + \max_{\alpha_{t+1,f} \in [\alpha_{t,f}, 1]} \left\{ -c_a \cdot y_t \cdot (\alpha_{t+1,f} - \alpha_{t,f}) + \frac{1}{1+r} \mathbb{E}\left[V_{t+1,f} | z_{t,f} \right] \right\} \right\} dG(c_o). \end{aligned}$$

E4. Entry decisions maximize the value function of entrants

$$V^e_{tf} = \int \max \left\{ 0, -c_o \cdot y_t + \max_{\alpha_{t+1,f} \in [\bar{\alpha}_t, 1]} \left\{ -c_a \cdot y_t \cdot (\alpha_{t+1,f} - \bar{\alpha}_t) + \frac{1}{1+r} \mathbb{E}\left[V_{t+1,f} | z_{tf} = z \right] \right\} \right\} dG(c_o),$$

where z denotes an entrant's productivity, and $\bar{\alpha}_t \equiv (\int_f \alpha_{tf} \cdot df)/(\int_f df)$.

E5. Starting from a distribution $\{\alpha_{0f}, z_{0f}\}$, the evolution of $\{\alpha_{tf}, z_{tf}\}$ is governed by the exogenous process for z, the endogenous process for α, and optimal entry and exit decisions.

Dynamics of task substitution and the labor share: We now explore the effects of changes in capital prices across tasks, which we model by a permanent increase in $q_t(x)$. This type of investment-specific technical change is in line with the work of Greenwood, Hercowitz and Krusell (1997), but we also consider the possibility that capital prices decline in some tasks more than in others. To save on notation, we focus on an equilibrium where firms produce all tasks in $[0, \alpha_{tf}]$ with capital. This will be the relevant scenario following

[11] In Appendix C.1, we demonstrate that our findings are not sensitive to different timing assumptions. In particular, we find similar results if firms observe $z_{t+1,f}$ before choosing $\alpha_{t+1,f}$.

a reduction in the cost of capital.

The task production function implies that the unit cost for a firm f at time t is

$$c_{tf} = \frac{1}{z_{tf}} \cdot \left(\Psi_t^k(\alpha_{tf}) + \Psi^\ell(\alpha_{tf}) \cdot w_t^{1-\eta} \right)^{\frac{1}{1-\eta}}. \tag{1}$$

This is the usual CES price index, with the difference that the *share parameters* $\Psi_t^k(\alpha_{tf})$ and $\Psi^\ell(\alpha_{tf})$ are now endogenous and depend on the mass of tasks that are automated:

$$\Psi_t^k(\alpha_{tf}) = \int_0^{\alpha_{tf}} (q_t(x) \cdot \psi^k(x))^{\eta-1} \cdot dx, \qquad \Psi^\ell(\alpha_{tf}) = \int_{\alpha_{tf}}^1 \psi^\ell(x)^{\eta-1} \cdot dx.$$

The share of capital in cost for a firm ε_{tf}^k—which equals the output-to-capital elasticity—and the share of labor in cost ε_{tf}^ℓ—which equals the output-to-labor elasticity—are then

$$\varepsilon_{tf}^k = \frac{\Psi_t^k(\alpha_{tf})}{\Psi_t^k(\alpha_{tf}) + \Psi^\ell(\alpha_{tf}) \cdot w_t^{1-\eta}}, \qquad \varepsilon_{tf}^\ell = \frac{\Psi^\ell(\alpha_{tf}) \cdot w_t^{1-\eta}}{\Psi_t^k(\alpha_{tf}) + \Psi^\ell(\alpha_{tf}) \cdot w_t^{1-\eta}},$$

and the labor share *in value added* is $s_{tf}^\ell = \varepsilon_{tf}^\ell / \mu$. In our model with CES demand, markups are fixed at μ, and the labor share is entirely driven by substitution decisions—captured by α_{tf}—and factor prices—captured by w_t and $q_t(x)$. These equations show that, as firms automate more of their tasks, they will increase the share of capital in both cost and value added, and reduce their labor shares.

To understand these dynamics, we first provide a lemma characterizing firms' automation decisions, and we then turn to studying how these decisions affect labor shares across firms. Let α_t^* denote the level of automation that would minimize firms' marginal cost of production. 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:

$$\frac{\psi^\ell(\alpha_t^*)}{q_t(\alpha_t^*) \cdot \psi^k(\alpha_t^*)} = w_t. \tag{2}$$

The following lemma characterizes automation decisions. In line with the existence of a fixed costs of automation per task, more productive firms automate more of their tasks.

Lemma 1 *Suppose that $\alpha_{tf} < \alpha_{t+1}^*$. Optimal automation decisions are given by $\alpha_{t+1,f} = \tilde{\alpha}_t(\alpha_{tf}, z_{tf})$, where $\tilde{\alpha}_t(\alpha_{tf}, z)$ is an increasing function of z that satisfies*

$$\lim_{z \to 0} \tilde{\alpha}_t(\alpha_{tf}, z) = \alpha_{tf}, \qquad \lim_{z \to \infty} \tilde{\alpha}_t(\alpha_{tf}, z) = \alpha_{t+1}^*.$$

The lemma shows that automation increases with firm size, exhibits history dependence, and is episodic. Firms that are highly productive and large will choose an automation level $\alpha_{t+1,f}$ close to α^*_{t+1}. Moreover, firms go through episodes of automation when z increases and $\tilde{\alpha}_t(\alpha_{tf}, z)$ exceeds α_{tf}.

To understand how changes in capital prices and the ensuing automation decisions by firms affect factor shares, we define two distinct elasticities of substitution. On the one hand, we have the elasticity of substitution between capital and labor holding the level of automation constant. In our model, this coincides with the *elasticity of substitution across tasks* η. On the other hand, we have the *induced elasticity of substitution*:

$$\eta^*_t = \eta + \frac{\partial \ln \Psi^k_t(\alpha)/\Psi^\ell(\alpha)}{\partial \ln \alpha} \Bigg/ \frac{\partial \ln \psi^\ell(\alpha)/(q_t(\alpha)\cdot\psi^k(\alpha))}{\partial \ln \alpha}.$$

The induced elasticity accounts for substitution across tasks (given by η) and the endogenous shifts in α^*_t in response to factor prices (the second term). Because optimal automation decisions are increasing in the wage and the level of $q(x)$ (from 2), this second term is always non-negative and the induced elasticity exceeds η. We let η^* denote the steady-state value of the induced elasticity.

We now provide two propositions that characterize the steady-state response of the economy to a permanent decline in capital prices.

Proposition 1 *Suppose that the economy is in a steady state with all firms having a common α^*, and let ε^ℓ and ε^k denote the common cost shares in this steady state. Following a permanent and uniform increase in $q(x)$ by $d\ln q(x) = d\ln q > 0$, the economy converges to a new steady state with wages rising by $d\ln w = (\varepsilon^k/\varepsilon^\ell)\cdot d\ln q > 0$, and automation rising by $d\ln\alpha^* > 0$. The aggregate share of labor in cost (or in value added) changes by*

$$d\ln\varepsilon^\ell = \frac{\varepsilon^k}{\varepsilon^\ell}\cdot(1-\eta^*)\cdot d\ln q.$$

Along the transition, firms automate $d\ln\alpha_{tf} \in [0, d\ln\alpha^*)$ tasks, depending on entry date and history of productivity draws. As a result, their labor share changes by

$$d\ln\varepsilon^\ell_{tf} = \frac{\varepsilon^k}{\varepsilon^\ell}\cdot\left[(1-\eta)-(\eta^*-\eta)\cdot\frac{d\ln\alpha_{tf}}{d\ln\alpha^*}\right]\cdot d\ln q$$

The case with $\eta < 1 < \eta^*$ will be particularly relevant for our analysis. The proposition shows that the aggregate labor share is controlled by the induced elasticity of substitution

η^*, and that it will decline as capital becomes cheaper.[12] The second part of the proposition shows that, even though the aggregate labor share declines, lower capital prices generate dispersion in changes in the labor share along the transition. At one extreme, we will have incumbents that do not reach a big enough scale to justify investments in automating additional tasks and will keep their α fixed at its initial value. For these firms, capital and labor are complements (since $\eta < 1$) and their labor share rises as capital prices drop and wages increase. In the other extreme, we have firms that receive positive productivity shocks and reach a sufficient scale to justify automating all the way up to α^*. For these firms, capital and labor become substitutes and their labor share falls as capital prices drop and wages increase. The median firm could see an increase in its labor share if the fixed cost of automation is enough to prevent it from automating a large fraction of tasks.

Although the case with $\eta < 1 < \eta^*$ is intuitive and some evidence supports it as a starting point, our model can generate the dynamics that we see in the data when $\eta < \eta^* \leq 1$. In this case, the labor share declines following improvements in technology that reduce the cost of capital used at marginal tasks (those around α^*), instead of uniformly at all tasks.

Proposition 2 *Suppose that the economy is in a steady state with all firms having a common α^*, and let ε^ℓ and ε^k denote the common cost shares in this steady state. Following a permanent increase in $q(x)$ for all $x > \alpha^*$ by $d\ln q(x) = d\ln q > 0$, the economy converges to a new steady state with wages rising by $d\ln w > 0$, and automation rising by $d\ln\alpha^* > 0$. The aggregate share of labor in costs (or value added) changes by*

$$d\ln\varepsilon^\ell = \varepsilon^k \cdot (1-\eta)\cdot d\ln w - \varepsilon^k \cdot (\eta^* - \eta)\cdot (d\ln q + d\ln w),$$

which is negative for small $d\ln q$. Along the transition, firms automate $d\ln\alpha_{tf} \in [0, d\ln\alpha^*]$ tasks, depending on entry date and history of productivity draws. As a result, their labor share changes by

$$d\ln\varepsilon^\ell_{tf} = \varepsilon^k \cdot (1-\eta)\cdot d\ln w_t - \varepsilon^k \cdot (\eta^* - \eta)\cdot \frac{d\ln\alpha_{tf}}{d\ln\alpha^*}\cdot (d\ln q + d\ln w_t)$$

To understand the difference with Proposition 1, consider the effects of a general change in capital prices across tasks on the labor share, and focus on the case with $\eta < 1$, which is the empirically relevant case. Lower capital prices have two effects. On the one hand, the reduction in capital prices for tasks above α^* leads to the substitution of capital for

[12] In steady state, the aggregate elasticity of substitution is exactly η^* due to diffusion. But even without diffusion, the aggregate elasticity of substitution will be close to η^* because large firms will automate almost all the way up to α^*, as we show in more detail in our numerical analysis.

labor in some of these tasks, which *always* reduces the labor share. On the other hand, the reduction in capital prices for tasks below α^* reduces the price of these tasks. This price effect lowers the share of these tasks in value added and *raises* the labor share.

A uniform decrease in the price of capital triggers both effects. As shown in Proposition 1, the substitution effect dominates on aggregate and for firms that automate their tasks up to the new optimal level if $\eta^* > 1$; whereas the price effect on the labor share is positive and dominates for incumbents that do not automate.

On the other hand, a localized decline in capital prices above α^* triggers only the substitution effect, which lowers the aggregate labor share and the labor share of firms that automate their tasks up to the new optimal level. This is partly counteracted by a positive price effect generated by higher wages—the term $(1-\varepsilon^\ell)\cdot(1-\eta)\cdot d\ln w$ in the equation for $d\ln\varepsilon^\ell$ in Proposition 2, and which captures the fact that higher wages increase the price of tasks produced by labor, raising their share in value added. However, the proof of the proposition shows that this effect is second order and dominated by the substitution effect. At the same time, smaller incumbents that do not automate additional tasks will see their labor shares increasing due to this positive price effect coming from higher wages.

Propositions 1 and 2 show that lower capital prices can drive the labor share decline independently of whether the elasticity of substitution is above or below 1. Through the lens of a task model, we have one of two interpretations. Either the price of capital declines uniformly across tasks and $\eta^* > 1$, as emphasized in Karabarbounis and Neiman (2013); Hubmer (2020) and in Proposition 1. Or this decline is more pronounced for new types of capital used at tasks where labor had a comparative advantage, as emphasized in Proposition 2. In this last case we could have $\eta^* \leq 1$, so that an econometrician exploiting variation in wages would estimate an elasticity of substitution below 1, as in Oberfield and Raval (2014). In both cases, the labor share of a typical firm that does not automate would increase along the transition due to price effects (an increase in the share of tasks performed by labor in value added, either due to lower capital prices at tasks initially produced with capital or higher wages at tasks initially produced with labor).

1.2 Calibration and quantitative results

This subsection explores the implications of Propositions 1 and 2 numerically, and shows that a calibrated version of our model reproduces the firm-level dynamics found in the US manufacturing sector. We study the effects of a uniform decline in the price of capital in the main text, and leave an analysis of localized changes to Appendix C.3. In our analysis, we take the economy in 1982 to be in a steady state with capital prices normalized to

$q_0(x) = 1$ and where all firms had the same level of automation α_0. We then explore the implications of a reduction in capital prices over time—an increase in $q_t(x)$. 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 post 1982 period also coincided with a pronounced decline in the price of equipment and software, which we identify as the main driving force behind the manufacturing labor share decline.[13]

Calibration: We parametrize capital and labor productivity across tasks as:

$$\psi^\ell(x) = A_\ell \cdot \left(x^{\frac{1-\eta-\gamma}{\gamma}} - 1\right)^{\frac{1}{1-\eta-\gamma}}, \qquad \psi^k(x) = A_k,$$

where A_ℓ and A_k denote standard factor-augmenting terms. With this specification, the production function of a firm f that automates all tasks up to α_f and rents k_f units of capital and ℓ_f units of labor becomes

$$(3) \quad y_f = z_f \cdot \left(\alpha_f^{\frac{1}{\eta}} \cdot (A_k \cdot k_f)^{\frac{\eta-1}{\eta}} + g(\alpha_f)^{\frac{1}{\eta}} \cdot (A_\ell \cdot \ell_f)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{1-\eta}}, \text{ with } g(\alpha_f) = \left(1 - \alpha_f^{\frac{\eta+\gamma-1}{\gamma}}\right)^{\frac{\gamma}{\eta+\gamma-1}}.$$

As we show in Appendix A, the induced elasticity of substitution is constant and equal to $\eta^* = \eta + \gamma$, where $\gamma > 0$ is an inverse measure of the strength of the comparative advantage of labor at higher-index tasks.[14] In what follows, we set $\eta = 0.5$ to match the estimate of the task-level elasticity of substitution in Humlum (2019) and $\gamma = 0.95$ 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 $\eta^* = 1.45$ (see Karabarbounis and Neiman, 2013; Hubmer, 2020).[15] Lastly, we calibrate A_k and A_ℓ to match the manufacturing labor share in 1982 and hourly wages in the sector.

We calibrate the demand system, the productivity process, and operating costs to match

[13] Appendix E shows that this is also the case for most economic sectors. After having labor shares that were stable for the 1947–1982 period, some sectors experienced a decline in their labor share of value added after this period. The 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).

[14] With this specification, we can think of firms as operating a standard CES production function in capital and labor and then paying a fixed cost to increase the CES share of capital. Equation (3) also relates task models and Putty-Clay models on the one hand, and models of production techniques a-la Caselli and Coleman (2006), on the other. For example, in Putty-Clay models, one can think of vintages of capital indexed by α as having different labor requirements captured by the CES shares α and $g(\alpha)$ in equation (3). In models a-la Caselli and Coleman (2006), one can think of α as indexing the choice over different production techniques, and $g(\alpha)$ as capturing the shape of the menu of techniques available.

[15] A reduction of capital prices by $d \ln q$ changes the labor share in costs by $\varepsilon^k \cdot (\eta^* - 1) \cdot d \ln q$. Hubmer (2020) documents that a reduction of capital prices of 1% lowers the labor share by about 0.10%. When the average markup is 1.15, we get $\varepsilon^k = 0.23$, and the estimates in Hubmer (2020) imply $\eta^* = 1.45$.

various moments from the US manufacturing sector. Panel II of Table 1 lists the calibrated parameters and moments targeted. We set the demand elasticity to $\sigma = 7.67$, which generates a common markup of 1.15. Turning to the process for productivities and firm dynamics, we assume that firm productivity z_f follows an AR1 process in logs:

$$\ln z_{t+1,f} = \rho_z \cdot \ln z_{t,f} + \varepsilon_{t+1,f},$$

where $\rho_z \in (0,1)$, $\varepsilon_{t+1,f} \sim N(\mu_z, \sigma_z)$, and $\mu_z = -\frac{(1-\rho_z)\cdot\sigma_z^2}{2\cdot(1-\rho_z^2)}$ so that the long-run mean of z_{tf} is normalized to one. We pick the dispersion of firm productivity to match the share of sales among the top 4 firms within 4-digit manufacturing industries reported by Autor et al. (2020), which corresponds to the top 1.1% of firms in each industry. We calibrate the fixed cost of operation and entry (as well as their dispersion) to match entry rates, exit rates, and the relative size of exiters and entrants reported in Lee and Mukoyama (2015) for US manufacturing.[16] Finally, we set the persistence of productivity ρ_z to 0.95, which we obtained from our estimates for the persistence of revenue TFP for manufacturing firms in Section 3. This estimate for the persistence of productivity is in the range of estimates for various TFP measures using US Census data in Foster, Haltiwanger and Syverson (2008).

The decline in the price of capital and the fixed cost of automating tasks: We now explore the adjustment of the economy following a decrease in the price of capital. We treat the economy in 1982 as being in steady state and calibrate the decline in capital prices required to match the observed decline in the manufacturing labor share from 1982 to 2012. In addition, we calibrate the fixed cost of automating tasks required to match the uneven adoption of automation technologies observed in the data by the end of our sample period. This approach assumes that the decline in the manufacturing labor share is entirely due to declining capital prices, and simply tries to recover the price decline needed to match it and inspects the implications of this driving force. Because of this, the results in this section should ve viewed as *possibility results*; they show that it is possible to have a coherent description of the manufacturing labor share decline driven by capital–labor substitution that fits the relevant facts.[17]

[16]Following Clementi and Palazzo (2016), we impose a Pareto distribution for the operating and entry cost with scale parameter $\underline{c}_o$ and tail coefficient ξ_o to match the frequency and relative size of exiters. We also match the relative size of entrants by modeling the entrant distribution as a log-normal that differs from the long-run distribution of z only insofar as it has a lower mean $\mu_e < 1$.

[17]An alternative approach would be to use direct measures of the shock. For example, one could use the observed decline in capital prices, as in Greenwood, Hercowitz and Krusell (1997). As we show below, this would yield similar results. Moreover, we prefer our shock calibration approach for two reasons. First, we only observe the average decline in the price of different capital goods. But these averages do not necessarily

Table 1: Calibration of the CES demand model for manufacturing

	Parameter		Moment	Data	Model
	I. Parameters related to the production function				
η	Task substitution elasticity	0.5	From Humlum (2019)	0.5	0.5
γ	Comparative advantage	0.95	Long-run K–L elasticity from Hubmer (2020)	1.45	1.45
	II. Parameters governing firm dynamics and productivities				
σ	Demand elasticity	7.67	Aggregate markup from Barkai (2020)	1.15	1.15
$\underline{c}_o$	Scale operating cost	$4.0 \cdot 10^{-7}$	Entry (=exit) rate from Lee and Mukoyama (2015)	0.062	0.063
ξ_o	Tail index operating cost	0.250	Relative exiter size from Lee and Mukoyama (2015)	0.490	0.490
μ_e	Entrant productivity	0.905	Relative entrant size from Lee and Mukoyama (2015)	0.600	0.600
σ_z	Std. dev. of $\ln z$ innovations	0.105	Top 4 firms' sales share in 1982 from Autor et al. (2020)	40.0%	40.0%
ρ_z	Productivity persistence	0.95	Revenue TFP persistence among manufacturing firms (see Section 3)		

Notes: 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 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 parameters in Panel II are jointly calibrated to match the corresponding moments.

In particular, we let $q_t(x) = q_t$ where q_t increases gradually from 1 in 1982 to a higher level q by 2012. We then calibrate q and c_a to match: (i) the 20 pp reduction in the manufacturing labor share between 1982 and 2012; and (ii) the fact that, by the end of our sample, firms in the top percentile of the employment distribution were 1.71 times more likely to have adopted new automation technologies than firms between the 50th and 75th percentile (see Acemoglu et al., 2021). In particular, we calibrate c_a so that

$$\frac{\mathbb{E}[\Delta\alpha_{tf}|\text{firm } f \text{ in employment P99+}]}{\mathbb{E}[\Delta\alpha_{tf}|\text{firm } f \text{ in employment P50–75}]} = 1.71,$$

where we take the increase in α_{tf} relative to the 1982 steady state as a measure of the adoption of new automation technologies.

map to tasks, nor imply that there has been a uniform decline in capital prices across tasks. The observed decline in capital prices can be driven by differential changes in $q_t(x)$ across tasks, with distinct implications for aggregates and factor shares. Second, when we turn to our model with endogenous markups, we won't have direct measures of rising competition or market size. In that case, calibrating both shocks puts both explanations on equal footing. We believe this is more compelling than treating capital prices as observed and rising concentration as a residual.

Panel I in Table 2 reports our calibrated shock and fixed cost of automation. We find that a decrease in capital prices of $d \ln q = 1.68$ between 1982 and 2012 and a fixed cost of automating tasks of $c_a = 0.354$ match (i) and (ii). The decline in the price of capital of 168 log points over the period 1982–2012 is also in the ball park of what we observe for computer-powered equipment and 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 goods prices). Moreover, the fixed cost of automation required to rationalize the data is of a reasonable magnitude. In response to lower capital prices in our model, aggregate spending on automation fixed costs as a share of manufacturing output rises from 0 in 1982 to a peak of 1% 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, this series provides an upper bound on expenditures required to develop and integrate the software and equipment required for automation. In line with this view, the model-implied spending on automation fixed costs amounts to half of the R&D share in the data, and the rise in spending on fixed costs aligns with the increasing share of R&D observed since the 80s.

Implications of lower capital prices: We now explore the implications of the inferred decline in capital prices for firms' labor shares and outcomes of interest. The results in Column 2 of Panel III show that our model matches the large decline in the manufacturing labor share at the same time as the median labor share remains unchanged and the unweighted mean labor share among manufacturing firms decreases by 2.3 pp— almost the same as in the data, where it decreased by 1.7 percentage points. Thus, the observation that the labor share has not fallen for the 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 due to the fixed cost required to automate new tasks.

The left panel in Figure 2 depicts the dynamics of the labor share across the distribution of firm productivity z_{tf}. The lines trace the labor share for firms at each percentile of the distribution at various points in time. In 1982, all firms have the same labor share independently of their size, since all operate technologies with the same level of automation. As capital prices decline, we see a clockwise rotation of this curve, with the labor share rising at the middle and the bottom of the firm-productivity distribution, but decreasing at the top, in line with Proposition 1.

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

		Model		
	Data	Benchmark	No fixed cost of automation	No diffusion
	(1)	(2)	(3)	(4)
	I. Parameters and inferred aggregate shocks			
$d \ln q$	.	1.68	1.68	1.65
c_a	.	0.354	0	0.241
	II. Targeted moments, 1982–2012			
Δ aggregate labor share	−0.20	−0.20	−0.23	−0.20
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1	1.72
	III. Typical firm labor share from Kehrig and Vincent (2020), 1982–2012			
Δ median labor share	0.030	-0.003	−0.228	0.070
Δ unweighted mean	-0.017	-0.023	−0.228	0.003
	IV. Other moments, 1982–2012			
Δ log top 4 firms' sales share	0.140	0.071	−0.030	0.053
Δ log top 20 firms' sales share	0.072	0.070	−0.025	0.055
Δ log productivity dispersion	0.050	0.059	-0.001	0.093
	V. Melitz–Polanec decomposition from Autor et al. (2020)			
Δ aggregate labor share	−0.185	−0.201	−0.228	−0.202
Δ unweighted incumbent mean	−0.002	0.006	−0.228	−0.032
Exit	−0.055	−0.004	0	−0.003
Entry	0.059	0.006	0	0.003
Covariance term	−0.187	−0.209	0	−0.171
	VI. Covariance decomposition from Kehrig and Vincent (2020)			
Market share dynamics	0.047	0	0	0
Labor share by size dynamics	−0.043	−0.111	−0.228	−0.120
Cross–cross dynamics	−0.232	−0.095	0	−0.084

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 labor share and the relative adoption of automation technologies by firm size (from Acemoglu et al., 2021). Column (3) displays a counterfactual economy with no fixed cost of automation. Column (4) displays a re-calibrated economy with no diffusion through entry. The change in industry concentration in Panel IV 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 V 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 VI reproduces the covariance decomposition from Kehrig and Vincent (2020, Figure 5), conducted for a balanced sample of firms and one long change 1982–2012.

The right panel plots the sales share by percentile of the productivity distribution relative to firms' sales shares in 1982. We see more productive firms increasing their market share along the transition, since these firms respond to the lower capital prices by automating more of their tasks and reducing their unit costs.[18] As shown in Panel IV of Table 2,

[18] Because of the assumed diffusion of automation technologies, these trends eventually revert over time. Without diffusion, the divergence in labor shares and the increase in sales dispersion are permanent.

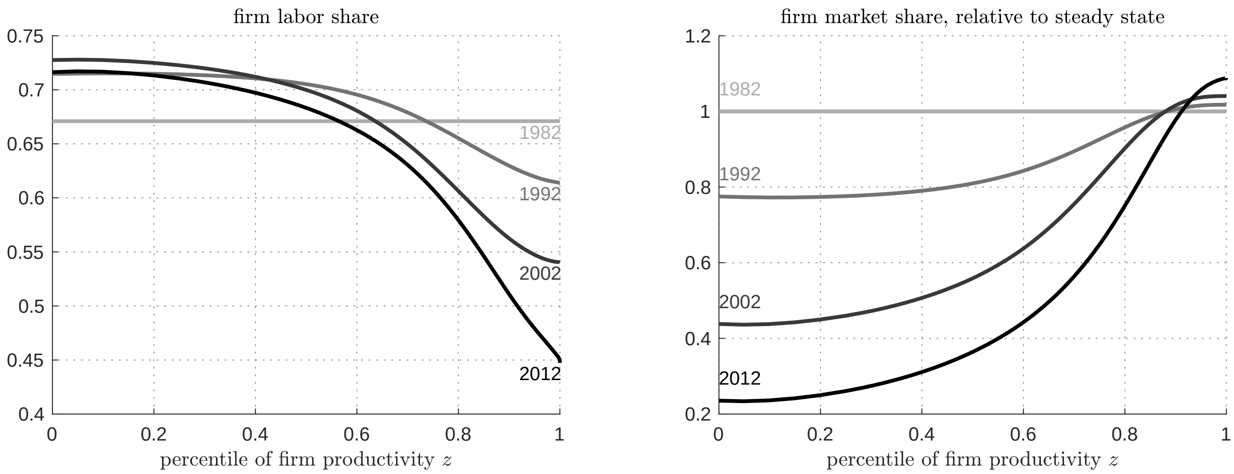

FIGURE 2: CROSS-SECTIONAL LABOR AND MARKET SHARES OVER THE TRANSITION. Firm labor shares in value added and firm sales shares by firm productivity z_{tf} from the benchmark model with CES demand. Sales shares, on the right, displayed relative to a percentile's sales share in the initial steady state in 1982.

by 2012, the uneven adoption of automation technologies in our model leads to a 7.1 log points (2.9 pp) increase in the share of sales among the top 1.1% firms, and a 7 log points (5 pp) increase in the share of sales among the top 5.5% firms in manufacturing. Empirically, Autor et al. (2020) document increases of 14 log points for the share of sales by the top 4 firms in each manufacturing industry (corresponding to the top 1.1% in our model) and 7.2 log points for the share of sales by the top 20 firms (the top 5.5% in our model). Thus, our model endogenously accounts for 50–95% of the observed increase in sales concentration in manufacturing. Moreover, our model provides an alternative explanation for the correlation between lower labor shares and higher sales concentration observed across US industries (see for example Barkai, 2020; Autor et al., 2020).[19]

To further investigate the predictions of our model for firms' labor share dynamics, we follow Autor et al. (2020) and decompose the decline in the manufacturing labor share using a Melitz–Polanec decomposition:

$$\begin{aligned}
\Delta s_t^\ell = \Delta \bar{s}_t^\ell & \qquad \text{(Change in unweighted incumbents' mean)} \\
+ \omega_{tX} \cdot \left(s_{tS}^\ell - s_{tX}^\ell\right) + \omega_{t'E} \cdot \left(s_{t'E}^\ell - s_{t'S}^\ell\right) & \qquad \text{(Contribution of exit and entry)} \\
+ \Delta \sum_f \left(\omega_{tf} - \bar{\omega}_t\right) \cdot \left(s_{tf}^\ell - \bar{s}_t^\ell\right) & \qquad \text{(Change in covariance).}
\end{aligned}$$

Here, Δs_t^ℓ denotes the change in the manufacturing labor share between two periods, t

[19] Comparing the 2000s to the 1980s, uneven automation in our model generates an increase in productivity dispersion of 5.9 log points. This matches the evidence in Decker et al. (2020), who estimate an increase in TFP dispersion of 5 log points for the US manufacturing sector over this period. Thus, our model leaves little room for other forces leading to higher productivity dispersion in manufacturing.

and t'. This can be decomposed into the change in the unweighted mean of labor shares among continuing firms, $\Delta \bar{s}^{\ell}_{t}$; 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 between their share of value added, ω_{tf}, and their labor share, s^{ℓ}_{tf}. The contribution of firms that exit the market is given by their share of value added in the baseline period, ω_{tX}, multiplied by the difference in the average labor share of continuing firms, s^{ℓ}_{tS}, and firms that exit, s^{ℓ}_{tX}. The contribution of firms that enter the market is given by their share of value added in the end period, $\omega_{t'E}$, multiplied by the difference in the average labor share of firms that enter, $s^{\ell}_{t'E}$, and continuing firms, $s^{\ell}_{t'S}$.

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 V in Table 2 reproduces Autor et al.'s manufacturing data and reports the decomposition from our model. In the data and model, the covariance term fully accounts for the aggregate decline in the labor share, with exit and entry and the change in the unweighted mean of incumbents' labor shares playing minor roles.[20] These results 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). The Melitz–Polanec decomposition does not 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 accounting sense, changes in the covariance term can be decomposed as

$$\begin{aligned}
\Delta \sum_f (\omega_{tf} - \bar{\omega}_t) \cdot (s^{\ell}_{tf} - \bar{s}^{\ell}_t) = & \sum_f \Delta(\omega_{tf} - \bar{\omega}_t) \cdot (s^{\ell}_{tf} - \bar{s}^{\ell}_t) && \text{(market share dynamics)} \\
& + \sum_f (\omega_{tf} - \bar{\omega}_t) \cdot \Delta(s^{\ell}_{tf} - \bar{s}^{\ell}_t) && \text{(labor share by size dynamics)} \\
& + \sum_f \Delta(\omega_{tf} - \bar{\omega}_t) \cdot \Delta(s^{\ell}_{tf} - \bar{s}^{\ell}_t) && \text{(cross-cross dynamics).}
\end{aligned}$$

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

[20]Exit and entry exhibit the same qualitative patterns as in the data, with both entering and exiting firms having labor shares that are higher than those of incumbents. However, these differences are not as pronounced as in the data, where many entering or exiting firms have labor shares that exceed 1, reflecting other elements of the life cycle of firms that are not in our model.

same time (the "cross-cross dynamics" term).

Kehrig and Vincent (2020) provide evidence suggesting that, in manufacturing, the cross–cross dynamics drove the decline in the labor share. Using a balanced sample of firms for 1982–2012, they document that the cross–cross dynamics contributed -23.2 pp to the decline in the manufacturing labor share, the labor share dynamics by size account for a 4.3 pp decline, and the market share dynamics actually increased the manufacturing labor share by 4.7 pp. Panel V in Table 2 provides the contributions of the three components above in the data and model.[21] Our model economy aligns with the data: in our model, the labor share dynamics by size contributed a 11.1 pp decline to the labor share, and the cross-cross dynamics contributed a 9.5 pp decline to the labor share (by construction, the market share dynamics do not affect the labor share in our model since all firms are assumed to have the same labor shares in 1982). The reason why the cross-cross dynamics are a key driver of the decline of the labor share in our model is that the firms that automate the most are those that receive a series of high productivity draws, allowing them to gain market share and recoup their fixed costs of automating tasks. Because these firms simultaneously gain market share and reduce their labor shares, this mechanism shows up as part of the covariance term in the Melitz–Polanec decomposition presented by Autor et al. (2020) and as part of the cross–cross term in the decomposition of Kehrig and Vincent (2020). We will see that our model with endogenous markups will do an even better job at matching the decomposition from Kehrig and Vincent (2020) quantitatively.

Fixed costs of adoption and diffusion through entry: For comparison, Column 3 in Table 2 summarizes the transitional dynamics of our model in response to the decline in capital prices calibrated above but now assuming that firms faced no fixed costs 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.

Finally, Column 4 shuts down the diffusion of automation technologies through entry and re-calibrates the decline in capital prices and the fixed cost of task automation required to match the labor share and technology adoption data. We see a more positive increase in the measures of a typical firm's labor share, but very similar dynamics and outcomes.

[21] We 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.

While introducing diffusion is useful for characterizing the steady state of our economy, it does not affect the dynamics of adjustment for the 1982–2012 period.

2 Capital–labor substitution vs. rising competition

This section extends our model to allow for differences in firms' markups. This 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.

2.1 Model and theoretical properties

Environment: We retain all elements of the model in section 1 but modify the demand system to allow for endogenous markups. We assume atomistic firms that face a demand derived from a Kimball aggregator (Kimball, 1995). Total output is defined implicitly by

$$\int_f \lambda_t \cdot H\left(\frac{y_{tf}}{\lambda_t \cdot y_t}\right) \cdot df = 1,$$

where H is an increasing and concave function. Normalizing the price of the final good to 1 yields the demand curve faced by a firm charging a price of p_{tf} as

$$y_{tf} = y_t \cdot \lambda_t \cdot D\left(\frac{p_{tf}}{\rho_t}\right),$$

where D is decreasing and given by the inverse function of $H'(x)$, ρ_t is an endogenous summary measure of competitors' prices, defined implicitly by

$$\int_f \lambda_t \cdot H\left(D\left(\frac{p_{tf}}{\rho_t}\right)\right) \cdot df = 1. \tag{4}$$

Finally, λ_t is an exogenous proxy for market size, which will serve to model increasing competition.[22]

We assume that demand satisfies Marshall's weak and strong second laws:[23]

(Marshall's weak second law) $\quad -x \cdot D'(x)/D(x)$ is greater than 1 and increasing in x

(Marshall's strong second law) $\quad x + D(x)/D'(x)$ is positive and log-concave

Marshall's weak second law requires that, as firms lower their prices, their demand becomes

[22]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.

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

more inelastic. This implies that firms with lower costs c_{tf} charge lower prices p_{tf} but higher markups μ_{tf}. Thus, larger firms will have higher markups. The strong second law requires marginal revenue to be positive and log-concave. This ensures that markups and prices, μ_{tf} and p_{tf}, are a log-convex function of costs, which implies lower passthroughs for more productive firms. The strong second law also implies that firm sales are a log-concave and decreasing function of costs.[24]

Equilibrium: Given a path for investment productivities $q_t(x)$, a path for market size λ_t, and an initial distribution of firms, an equilibrium is defined as before. The only difference is that now we also solve for the competitors' price index ρ_t using (4).

Effects of rising competition: As in our baseline model, the equilibrium of the economy converges to a steady state where $\alpha = \alpha^*$. This section characterizes the steady-state effects of rising competition, which we capture through an increase in λ_t. An increase in λ_t proxies for a rise in the effective market size faced by firms, which could be due to expanding trade or advances in information and communications technology that facilitate sales.[25]

Because in steady state all firms have the same level of automation, we only keep track of firms that differ in their productivity level z. Denote by μ_z the markup charged by a firm of productivity $z_f = z$, and by ω_z its sales share. Finally, let m_z denote the mass of firms of productivity z.

Proposition 3 *A permanent increase in λ_t has the following effects in the steady state distribution of firm markups and sales:*

- *μ_z decreases for all z;*

[24] We impose Marshall's second weak and strong law for three reasons. First, these assumptions receive support from the data (Baqaee and Farhi, 2020a). Second, this demand structure offers a tractable way of capturing the type of pricing dynamics in oligopolistic competition models, where large firms recognize that an increase in their price will have a disproportionate effect on the industry price index, making their revenue less responsive to changes in their price. Third, these assumptions ensure that more productive firms are larger but charge higher markups, and that an increase in competition will reallocate economic activity towards large firms with high markups. This is the key mechanism driving the decline of the labor share in theories that see rising competition as the main driving force behind the labor share dynamics. For example, Autor et al. (2020) use this model to illustrate how competition can reduce the labor share.

[25] Changes in λ_t can be interpreted in a reduced-form way as capturing all forces that generate a reallocation of sales towards the most productive firms in an industry *for a given* productivity distribution. This can include demand-side forces, but also supply-side forces such as improvements in advertisement or customer targeting and acquisition. However, changes in λ_t do not capture the role of other forces that generate sales concentration via rising productivity dispersion across firms, such as the uneven adoption of automation technologies considered here, the widening gaps between leaders and followers considered in Akcigit and Ates (2019) and Olmstead-Rumsey (2019), or the shifting age composition of firms in Hopenhayn, Neira and Singhania (2018).

- *for $z > z'$, $\mu_z/\mu_{z'}$ decreases;*
- *for $z > z'$, $\omega_z/\omega_{z'}$ increases.*

One can think of market size as increasing the demand faced by all firms, generating tougher competition for workers. The proposition shows that as rising competition for workers increases real wages and production costs, firms are pushed towards the more elastic segments of their demand curves and respond by reducing their markups. However, the reduction in markups is not uniform. Because large firms have smaller passthroughs, they respond via a modest increase in their prices (and a large reduction in their markup). Small firms on the other hand respond via a more sizable increase in their prices (and a smaller reduction in their markups). As a result, an increase in λ_t reallocates economic activity and labor towards the largest and most productive firms in the industry.

These responses by firms also generate an ambiguous contribution of markups to the aggregate labor share. On the one hand, firms of a given productivity level reduce their markups, which contributes to an increase in the aggregate labor share. On the other hand, the reallocation of economic activity from small firms with low markups to large firms with high markups contributes to a decline in the aggregate labor share.

PROPOSITION 4 *The aggregate labor share is $s^\ell = \varepsilon^\ell/\mu$, where the aggregate markup μ is a sales weighted harmonic mean of firm-level markups:*

$$\frac{1}{\mu} = \int_z \frac{1}{\mu_z} \cdot \omega_z \cdot m_z \cdot dz.$$

Holding the distribution of productivity m_z constant, an increase in λ increases the aggregate markup if the distribution of productivity is log-convex (i.e., more convex than Pareto), lowers it if the distribution of productivity is log-concave (i.e., less convex than Pareto), and leaves it unchanged if the distribution of productivity is log-linear (i.e., Pareto).[26]

The proposition shows that the effects of rising competition on the labor share through markups depend on the distribution of firm productivity. 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 relevant for understanding the calibrated effects of rising competition on the labor share and the aggregate markup.

[26] We refer to a distribution as log convex (log concave) if its PDF is log convex (log concave).

2.2 Calibration for manufacturing and retail

We now explore the transitional dynamics in response to rising competition and lower capital prices. We start by calibrating the model to the US manufacturing sector and then describe how we calibrate the model to the retail sector. We then explain how we use the model to calibrate the decline in the price of capital—captured by q, as in the previous section—and the increase in competition, λ, required to match the observed behavior of the labor share and sales concentration in each sector.

Manufacturing: As before, we calibrate the model under the assumption that manufacturing firms were in their steady state in 1982. We use the same parametrization of the production function and task productivities from the previous section, and summarized in Panel I of Table 3.

Relative to the previous section with a CES demand, we make two modifications. First, 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. With this specification, the demand elasticity faced by a firm with price p_{tf} is

$$\text{demand elasticity}(p_{tf}) = \sigma \cdot D\left(\frac{p_{tf}}{\rho_t}\right)^{-\frac{\nu}{\sigma}}, \tag{5}$$

which decreases as p_{tf} falls, so that more productive firms face more inelastic demand.[27] Here, σ controls the average demand elasticity faced by firms, and the *super-elasticity* ν/σ controls the extent to which markups rise for more productive firms and their lower passthroughs. (If $\nu = 0$, the demand system simplifies to the standard CES aggregator.) The super elasticity thus determines the net effect of rising competition on sales concentration.

Second, we adopt a new process for firm-level productivities z_{tf}. With a non-CES demand system, we can no longer assume a log-normal productivity distribution. As discussed above, under Marshall's second laws, sales are a log-concave function of costs (and hence productivity). Because the sales distribution is approximately Pareto, we need to entertain the possibility that the productivity distribution is more convex than Pareto. To do so, we assume that productivity is determined by a latent factor $\tilde{z}_{tf}$ that follows an AR(1) process as before and that determines productivity as:

$$z_{tf} = \exp\left(F^{-1}_{Weibull(n,\zeta)}\left(\Phi\left(\tilde{z}_{tf}\right)\right)\right), \quad \text{where} \quad \tilde{z}_{t+1,f} = \rho_z \cdot \tilde{z}_{tf} + \varepsilon_{tf}$$

[27] The full specification for H and the derivation of equation (5) are provided in Appendix B.

Here, Φ denotes the Gaussian cdf, and $F^{-1}_{Weibull(n,\zeta)}$ the inverse CDF of a Weibull random variable with shape parameter $n > 0$ and scale parameter $\zeta > 0$. The innovations are drawn from $\varepsilon_{tf} \sim N(\mu_z, \sigma_z)$, where μ_z and σ_z are normalized so that the long-run distribution of $\tilde{z}_{tf}$ is a standard normal.[28] This specification implies that $\ln z_{tf}$ follows 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 $\ln z_{tf}$ is log-linear, or equivalently, the limit distribution for z_{tf} is Pareto with tail index ζ. For $n < 1$, the density of $\ln z_{tf}$ is log-convex, or equivalently, the limit distribution for z_{tf} is more convex than Pareto. As shown in Proposition 4, n will be the key parameter determining the net effect of rising competition on the aggregate labor share of an industry.

We jointly calibrate the parameters governing markups $\{\sigma, \nu\}$ and the firm productivity distribution $\{\zeta, n\}$ to match the average manufacturing markup, the ratio of the (unweighted) mean firm to the aggregate labor share, and the share of sales among the top 4 and top 20 firms in manufacturing in 1982. In practice, these parameters are jointly calibrated with the fixed cost of operation and its dispersion and mean entrant productivity to match the same moments from Lee and Mukoyama (2015) introduced above, and after setting $\rho_z = 0.95$. Panel II in Table 3 summarizes the parameters and moments informing their calibration. Even though some of our parameters are calibrated to match the same moments as in the CES model, we obtain new estimates due to the different demand structure.

We calibrate $\sigma = 6.0$ to match an aggregate markup of 1.15. More importantly, we estimate a super-elasticity ν/σ of 0.22 to match a 1.1 ratio between the unweighted average of the labor share among manufacturing firms and the aggregate manufacturing labor share before 1982 from Kehrig and Vincent (2020). Because we assume that the economy is in a steady state initially, our calibration attributes all the difference in labor shares before 1982 to markups, which in turn reflect changes in the curvature of demand between the average firm and more productive firms. The log-concave demand system implies that smaller firms have lower markups; therefore, the unweighted mean labor share across manufacturing firms

[28] As before, we assume that entrants draw an initial latent factor $\tilde{z}_{tf}$ from a normal distribution that differs from the long-run distribution of $\tilde{z}$ only insofar as its mean is shifted to the left by $\ln \mu_e < 0$, which we calibrate to match the relative size of entrants in the data.

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

	Parameter		Moment	Data	Model
	I. Parameters related to production function				
η	Task substitution elasticity	0.5	From Humlum (2019)	0.5	0.5
γ	Comparative advantage	0.95	Long-run K–L elasticity from Hubmer (2020)	1.45	1.45
	II. Parameters governing firm dynamics and productivities in 1982 steady state				
ν/σ	Demand super-elasticity	0.22	Ratio of mean firm to aggregate labor share	1.10	1.09
σ	Demand elasticity	6.00	Aggregate markup	1.15	1.15
ζ	Weibull scale	0.077	Top 20 firms' sales share	69.7%	69.9%
n	Weibull shape	0.74	Top 4 firms' sales share	40.0%	40.0%
$\underline{c}_o$	Scale operating cost	$2 \cdot 10^{-7}$	Entry (=exit) rate	0.062	0.062
ξ_o	Tail index operating cost	0.24	Size of exiters	0.490	0.485
μ_e	Entrant productivity	0.882	Size of entrants	0.600	0.600
ρ_z	Productivity persistence	0.95	Revenue TFP persistence among manufacturing firms		

Notes: The ratio of the (unweighted) mean firm labor share to the aggregate manufacturing labor share is computed based on the replication data from Kehrig and Vincent (2020). 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 remaining data moments follow the model with CES demand, see Table 1. Fixing productivity persistence, the remaining seven parameters in Panel II are jointly calibrated to match the seven corresponding moments.

will exceed the aggregate labor share of the sector by a factor that increases in the super-elasticity of demand. A super elasticity of zero yields a ratio of 1, while a higher super elasticity of 0.22 matches the ratio of 1.1 observed in the data.[29]

For the productivity distribution, we calibrate $\{\zeta, n\}$ to match the share of sales by the top 4 firms and the top 20 firms within manufacturing industries, which roughly corresponds to the top 1.1% and top 5.5% share of sales. Intuitively, a higher top 5.5% sales share indicates a thicker tail of the productivity distribution—a higher ζ. Moreover, conditional on the top 5.5% share, a higher top 1.1% share requires a lower value of n, which indicates a more than proportional increase in productivity as we move to the top of the sales distribution. By targeting these two moments we find that $n = 0.74$—a small deviation from Pareto and thus a moderate degree of log-convexity—fits the manufacturing sales data. The fact that we estimate a log-convex distribution for productivity is in line with theory. Because sales are a log-concave function of productivity, this is needed to match

[29] Our calibrated demand super-elasticity (0.22) is close to the preferred estimate in Edmond, Midrigan and Xu (2018) of 0.16, who estimate this super-elasticity to match labor share dispersion by firm size in US Census data. Appendix C.4 provides a robustness exercise where we use a lower value of 0.16 for the super-elasticity of demand. This slightly weakens the contribution of the rising competition shock to the labor share decline, but overall the results do not change much.

the roughly log-linear distribution of sales observed in the data.

Retail: We follow the same calibration approach for retail and report our results in Table 4. Relative to manufacturing, the main difference for retail is that we calibrate a lower value of $n = 0.47$ (implying more log-convexity of the productivity distribution), which is necessary to match the high sales share of the top 4 firms in each 4-digit 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%. We also set the persistence of productivity ρ_z to 0.86, which matches our estimates for revenue TFP among retail firms in Section 3.[30]

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

	PARAMETER		MOMENT	DATA	MODEL
	I. Parameters related to production function				
η	Task substitution elasticity	0.5	From Humlum (2019)	0.5	0.5
γ	Comparative advantage	0.95	Long-run K–L elasticity from Hubmer (2020)	1.45	1.45
	II. Parameters governing firm dynamics and productivities in 1982 steady state				
ν/σ	Demand super-elasticity	0.22	Imputed from manufacturing		
σ	Demand elasticity	8.95	Aggregate markup	1.15	1.15
ζ	Weibull scale	0.0128	Top 20 firms' sales share	29.9%	29.9%
n	Weibull shape	0.47	Top 4 firms' sales share	15.1%	15.1%
$\underline{c}_o$	Scale operating cost	$4.6 \cdot 10^{-6}$	Entry (=exit) rate	0.062	0.062
ξ_o	Tail index operating cost	0.320	Size of exiters	0.490	0.494
μ_e	Entrant productivity	0.855	Size of entrants	0.600	0.600
ρ_z	Productivity persistence	0.86	Revenue TFP persistence among retail firms		

Notes: 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 remaining data moments follow the model with CES demand, see Table 1. Fixing productivity persistence and the demand super-elasticity, the remaining six parameters in Panel II are jointly calibrated to match the six corresponding moments.

2.3 Quantifying the role of competition and capital–labor substitution

We now use the model to calibrate the change in q and λ required to match the observed decrease in the aggregate labor share and the rise in concentration in manufacturing and

[30]For retail and all sectors outside of manufacturing, we lack consistent data on exit and entry rates, as well as the relative size of entrants and exiters. For this reason, we use the same moments that we targeted in our calibration for the manufacturing sector from Lee and Mukoyama (2015). Finally, we keep the same super-elasticity of demand of 0.22 and provide robustness checks to using different values in Appendix C.4.

retail from 1982 to 2012. Our exercise puts the capital–labor substitution and rising competition explanations on equal footing and uses the model to infer the contribution of these driving forces. 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, or capital prices, q. Thus, our exercise provides an answer to the question: if changes in the labor share and sales concentration are due to lower capital prices and rising competition, how important have each of these factors been in explaining the observed outcomes in manufacturing and retail?

Manufacturing: As before, we treat the economy in 1982 as being in steady state. We then calibrate the increase in q—lower capital prices—and λ—the rising competition shock—required to match (i) the observed decline in the manufacturing labor share and (ii) the increase in sales concentration. In addition, and as we did in the previous section, we calibrate the fixed cost of automation per task to match (iii) the higher adoption rate of automation technologies among large firms documented in Acemoglu et al. (2021).

Column 1 in Table 5 summarizes the manufacturing data and Column 2 reports our results. To match the trends in (i)–(ii), our model requires a decline in the price of capital of 156 log points and a mild increase in competition of 5%. The reason why we calibrate a small increase in λ is that the rise in sales concentration in this sector has been modest. Manufacturing is the sector with the lowest increase in the share of sales accruing to top firms from 1982–2012 according to the data in Table 1 of Autor et al. (2020). Moreover, this small increase in λ has a negligible effect on the manufacturing labor share, since the productivity distribution in this sector is close to Pareto. The model then requires a large decline in the price of capital to match the large decline in the manufacturing labor share.

In response to these two shocks, our model provides a good fit to the manufacturing data, matching the aggregate labor share decline, the rise in concentration, and the dynamics of the labor share across firms. As before, a small fixed cost of automating tasks $c_a = 0.29$, which we calibrated to match the patterns in Acemoglu et al. (2021), is enough to ensure that the (un-targeted) labor share of the typical manufacturing firm increases despite the lower capital prices. In particular, Panel III shows that the median firm labor share rises by 4.7 pp and the unweighted mean increases by 1.6 pp between 1982 and 2012.

Panels IV and V of Table 5 describe the dynamics of the labor share across firms using the Melitz–Polanec and the covariance decomposition introduced in Section 1.2. In line

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

		Model		
	Data	Benchmark	Only effects of $d\ln q$	Only effects of $d\ln\lambda$
	(1)	(2)	(3)	(4)
	I. Parameters and inferred aggregate shocks			
$d\ln q$		1.56	1.56	0
$d\ln\lambda$		0.05	0	0.05
c_a		0.29	0.29	0.29
	II. Targeted moments, 1982–2012			
Δ aggregate labor share	−0.199	−0.199	−0.184	-0.003
Δ log top 4 firms' sales share	0.140	0.140	0.079	0.058
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1.59	4.92
	III. Typical firm labor share and other moments, 1982–2012			
Δ median labor share	0.030	0.047	0.033	0.010
Δ unweighted mean	-0.017	0.016	0.006	0.008
Δ log top 20 firms' sales share	0.072	0.135	0.096	0.039
Δ log productivity dispersion	0.050	0.071	0.060	0.000
	IV. Melitz–Polanec decomposition from Autor et al. (2020)			
Δ aggregate labor share	−0.185	−0.199	−0.184	-0.003
Δ unweighted incumbent mean	−0.002	0.055	0.045	0.010
Exit	−0.055	−0.014	−0.014	−0.018
Entry	0.059	0.013	0.014	0.016
Covariance term	−0.187	−0.252	−0.228	−0.011
	V. Covariance decomposition from Kehrig and Vincent (2020)			
Market share dynamics	0.047	0.067	0.071	0.075
Labor share by size dynamics	−0.043	−0.025	−0.022	0.079
Cross-cross dynamics	−0.232	−0.239	−0.227	−0.155
	VI. Markups, 1982–2012			
Δ log aggregate markup	.	0.013	0.012	0.001
Within firm change in markup	.	−0.023	−0.018	−0.014
Reallocation to high-markup firms	.	0.036	0.030	0.015

Notes: Column (2) reports the findings from our benchmark model, which jointly calibrates (i) a uniform decline in the capital price (over 1982–2012), (ii) an increase in competition (over 1982–2012), and (iii) the automation fixed cost to replicate (i) the change in the aggregate manufacturing labor share (BLS/BEA integrated industry-level production account), (ii) the increase in the top 4 firms' sales share within 4-digit manufacturing industries (Autor et al., 2020, Table 1), and (iii) the relative adoption of automation technologies by firm size (from Acemoglu et al., 2021). Column (3) shows results when shutting down the competition shock, column (4) when shutting down instead the price of capital shock. 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.

with the data, we find a crucial role for a decline in the covariance between firm sales and their labor share in explaining the decline in the labor share. The decomposition in Panel V shows that this is fully explained by firms that expand at the same time as they reduce their labor shares (the cross-cross dynamics term), as in the data. The cross-cross dynamics term is now more negative because firms that expand not only automate more tasks when they do so, but also raise their markups. As in the data, our model now produces a positive market share effect. This is because firms that had low labor shares in 1982 were already large. Because of mean reversion in the productivity process, these firms will tend to loose market share over time, contributing to an increase in the labor share.[31]

To understand the contribution of each of these shocks, we provide counterfactual scenarios where we shut them down sequentially. In Column 3, we shut down the increase in competition. Capital–labor substitution explains 18.4 of the observed 19.9 pp decline in the manufacturing labor share. Moreover, capital–labor substitution explains 7.9 log points of the observed 14 log points increase in sales concentration among the top 4 firms in the sector. On the other hand, 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 an additional 5.8 log points.[32]

Panel VI of Table 5 summarizes the predictions of our model for markups. The labor share in an industry can be written as $s^\ell = \varepsilon^\ell/\mu$, where ε^ℓ is the share of labor in costs for the industry and μ is the aggregate 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 4, this is the relevant notion of an aggregate markup.[33] The decomposition shows that the manufacturing labor share

[31] Similar forces operate in a stationary equilibrium of our model. In particular, because of Marshall's weak second law, as firms cycle through high and low productivities, they will change their markups and go through transient cycles of low and high labor shares, similar to those document in Kehrig and Vincent (2020). These ergodic cycles generate a positive contribution of market share and labor share by size dynamics to the overall labor share behavior, and a negative contribution of the cross-cross term that exactly offsets them. A decline in capital prices changes the nature of these cycles, so that now, firms also automate during expansions, generating a permanent decline in their labor share, which generates a more negative cross-cross term and a negative contribution of the labor share by size dynamics that dominate the market share component, as in the data.

[32] In this counter-factual scenario, some firm will automate additional tasks in response to the higher equilibrium wages. This channel is not too large and only accounts for 0.2 pp of the labor share decline due to rising competition.

[33] This decomposition follows from the chain of identities

$$s^\ell = \frac{\sum_f s^\ell_f y_f}{\sum_f y_f} = \frac{\sum_f s^\ell_f y_f}{\sum_f \frac{1}{\mu_f} y_f} \frac{\sum_f \frac{1}{\mu_f} y_f}{\sum_f y_f} = \varepsilon^\ell \cdot \frac{1}{\mu}.$$

The last step uses 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 industry.

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 1.3% (from 1.15 to 1.165). However, as anticipated in Proposition 4, this net effect masks two opposing 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 reduced the manufacturing markup by 2.3% 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. The contribution of reallocation to markups is given by

$$\text{markup reallocation} = \Delta \ln \mu - \sum_f \omega_f \cdot \Delta \ln \mu_f,$$

which increased the manufacturing markup by 3.6% during this period. Interestingly, the results in columns (3) and (4) show that both lower capital prices and rising competition reallocate economic activity towards firms with large markups in manufacturing. The fact that rising competition leads to this form of reallocation is in line with Proposition 3. Lower capital prices have a similar effect because automation favors the expansion of large firms, which are precisely the ones with higher markups.[34] In sum, markups explain 0.9 of the 20 pp decline in the labor share of manufacturing, with the reallocation component explaining 2.1 pp (10%) of the decline.

Retail: We now calibrate the decline in capital prices and increase in competition required to match the observed labor share decline and the rising sales concentration in retail. Table 6 summarizes the retail data and presents the calibrated increase in q (lower capital prices) and λ (rising competition). As before, we calibrate the fixed cost of automating tasks jointly to match the higher adoption of automation technologies by large firms in the data.[35]

In retail, we observe a decline in the labor share of 12.7 pp as well as a rise in sales

[34] Baqaee and Farhi (2020b) show that shocks that reallocate economic activity towards firms with higher markups increase allocative efficiency and aggregate productivity. Proposition 1 in their paper shows that the contribution of changes in allocative efficiency to TFP are given by $\Delta \ln \mu - \sum_f \omega_f \Delta \ln \mu_f$—the markup reallocation term. Thus, our model predicts that improvements in allocative efficiency brought by rising competition and lower capital prices generated a 3.6% increase in manufacturing TFP from 1982 to 2012.

[35] Acemoglu et al. (2021) report similar relative adoption rates inside and outside of manufacturing. For this reason, we keep the same target as before.

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

	Data	Model: Benchmark	Model: Only effects of $d\ln q$	Model: Only effects of $d\ln\lambda$
	(1)	(2)	(3)	(4)
I. Parameters and inferred aggregate shocks				
$d\ln q$	.	0.79	0.79	0
$d\ln\lambda$	.	0.30	0	0.30
c_a	.	0.33	0.33	0.33
II. Targeted moments, 1982–2012				
Δ aggregate labor share	−0.127	−0.127	−0.051	−0.048
Δ log sales concentration	0.546	0.538	0.031	0.489
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1.34	2.19
III. Typical firm labor share and other moments, 1982–2012				
Δ median labor share	.	0.041	-0.007	0.042
Δ unweighted mean	.	0.029	-0.010	0.030
Δ log productivity dispersion	.	0.016	0.004	0.002
IV. Markups, 1982–2012				
Δ log aggregate markup	.	0.049	0.005	0.040
Within firm change in markups	.	−0.013	−0.010	−0.015
Reallocation to high-markup firms	.	0.062	0.015	0.055

Notes: Column (2) reports the findings from our benchmark model, which jointly calibrates (i) a uniform decline in the capital price (over 1982–2012), (ii) an increase in competition (over 1982–2012), and (iii) the automation fixed cost to replicate (i) the change in the aggregate retail labor share (BLS/BEA integrated industry-level production account), (ii) the average log change in the top 4 as well as top 20 firms' sales share within 4-digit retail industries (Autor et al., 2020, Table 1), and (iii) the relative adoption of automation technologies by firm size (from Acemoglu et al., 2021). 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.

concentration among the top 4 firms of 14.0 pp and a similar rise among the top 20 firms of 16.3 pp. Averaging over these two measures, this represents a vast increase in sales concentration of 55 log points (74%). For comparison, sales concentration increased by merely 14 log points in manufacturing. Our model can explain the observed patterns for retail with a smaller decline in the price of capital of 79 log points and an increase in competition of 30 log points—an order of magnitude larger than in manufacturing.[36]

The different inference obtained for manufacturing and retail is due to two key factors: (i) the vast increase in sales concentration in retail vis-a-vis the modest increase in manufacturing, and (ii) the more log-convex distribution of productivity in the retail sector. In retail, our model requires a large increase in competition to match the observed rise in

[36] In retail and for other sectors outside of manufacturing, we calibrate the shocks to match the average increase in concentration among the top 4 and top 20 firms in each 4-digit industry. We do this because the top 4 firms are a tiny fraction (0.023%) of all retailers, whereas they account for 1.1% of all manufacturing firms. Targeting instead the average top 4 and top 20 sales share increase in manufacturing as well (11 logs points) yields an even smaller role for the inferred λ-shock.

concentration. Moreover, because the productivity distribution in retail is more log convex than in manufacturing ($n = 0.47$ in retail vs. $n = 0.74$ in manufacturing), this increase in competition by itself has a more pronounced effect on the labor share, leaving a smaller role for lower capital prices.

Our model provides a good fit to the available data for retail and its evolution over time. As before, a small fixed cost of automating tasks is enough to match the adoption data and ensure a rise in the labor share of the typical retail firm. We find that the unweighted mean labor share rises by 1.6 pp and the median by 4.1 pp. Though we do not have direct data on these moments, this is in line with the evolution of payroll shares of sales reported in Autor et al. (2020), who find that the unweighted mean of payroll shares in retail increased by 4.4 pp between 1982 and 2012.

Columns 3 and 4 report the effects of the increase in q and λ separately. The decline in capital prices explains 40% of the decline in the retail labor share (−5.1 pp), but only 5% of the increase in sales concentration among the top 4 and top 20 firms within 4-digit retail industries. The increase in market size now explains close to 40% of the decline in the labor share (−4.8 pp) and about 90% of the increase in concentration.[37] The counterfactual scenarios in Columns 3 and 4 also point to a sizable interaction between these two shocks. The estimated q-shock by itself causes the aggregate labor share to decline by 5.1 pp, while the estimated λ-shock generates a 4.8 pp decline. Yet, in combination the two shocks generate a decline of 12.7 pp. Thus, their interaction accounts for 2.8 pp—about 20% of the total decline. Two mechanisms are responsible for this interaction. First, rising competition reallocates activity towards more automated firms with lower unit costs. This form of reallocation also contributes to the decline in the labor share. Second, rising competition implies that more productive firms account for a greater share of sales, generating extra incentives for automation in response to lower capital prices.

Turning to markups, our model predicts an increase in the aggregate markup in retail of 4.9% (from 1.15 to 1.21). This is the result of a 1.3% decrease in the within-firm component and a 6.2% increase driven by reallocation to high-markup firms. Most of the rise in markups is in this case explained by rising competition.[38] In sum, markups explain 3.5 of the 12.7 pp decline in the labor share of retail, with the reallocation component generating 4.5 pp (35%) of the decline.

[37] As before, part of the decline in the labor share arises due to firms automating more tasks in response to higher equilibrium wages. This effect alone generates a 2.0 pp decline in the labor share.

[38] Building on footnote 34, we find that, in retail, improvements in allocative efficiency due to rising competition contributed a 5.5% increase in sectoral TFP between 1982 and 2012; while lower capital prices had a smaller effect on allocative efficiency.

Other sectors: We also conducted our decomposition for other economic sectors, including wholesale, and utilities & transportation.[39] Figure 3 summarizes our findings and for reference contrasts them with our results for manufacturing and retail. In wholesale and utilities & transportation, the inferred q-shock accounts for less than one third of the labor share declines, while the inferred λ-shock accounts for two thirds of the sectoral labor share declines. On the other hand, the increase in competition generated by the λ-shock accounts for almost all the increase in sales concentration in both sectors.[40] We conclude that while the fall in the manufacturing labor share is almost entirely attributed to lower capital prices and the automation of additional tasks, in non-manufacturing sectors there is an important role for rising competition.

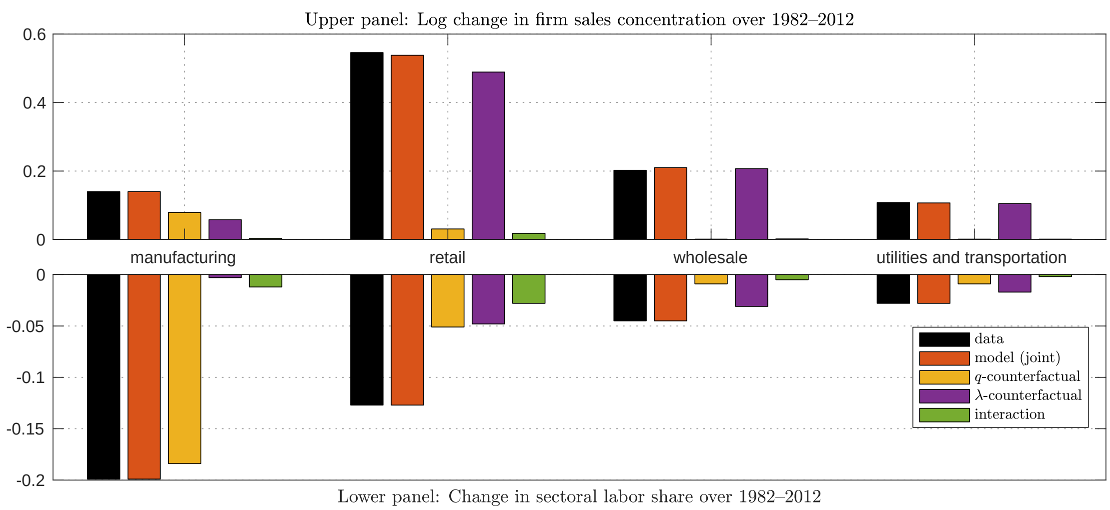

FIGURE 3: MODEL-BASED DECOMPOSITION OF LABOR SHARE AND SALES CONCENTRATION CHANGES. For each sector, the upper panel displays the log change in firm sales concentration (i) in the data (Autor et al., 2020, Table 1), (ii) in the benchmark model with q- and λ-shocks jointly calibrated, (iii) in a model counterfactual that keeps only the q-shock active, (iv) in a model counterfactual that keeps only the estimated λ-shock active; (v) displays the interaction term, defined as (ii - iii - iv). The lower panel shows sectoral labor share changes in data (BEA-BLS) and model. See Tables 5, 6 and 11 for details.

3 DIRECT EVIDENCE ON MARKUPS AND OUTPUT ELASTICITIES

Our quantitative exercise shows that most of the decline in the manufacturing labor share is due to lower capital prices and the ensuing automation of tasks, which increases large

[39]We relegate the calibration and data details to Appendix D. Also, we 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.

[40]Figure 3 also illustrates how the identification of the q- and λ-shock works in our model. In all sectors, the q-shock generates a decline in the labor share that is large relative to its effect on sales concentration. Instead, the λ-shock generates the opposite pattern, loading much more on the increase in concentration than the decline in the labor share.

firms output-to-capital elasticities. In retail and other sectors outside of manufacturing, the reallocation of economic activity to high-markup firms played a more prominent role, and explains half of the decline in the labor share. This section provides direct evidence on the behavior of markups and output-to-capital elasticities across firms. Markups and output elasticities are not directly observed, and so we rely on estimates for Compustat firms based on their revenue, expenditures in variable inputs, capital, and investment. Estimating production functions and markups using Compustat relies on strong assumptions. Our estimates in this section must be interpreted with the same caution required to approach previous empirical estimates of markups relying on these data and using similar methods.

3.1 Estimating output elasticities

Consider a firm that produces output by combining capital, k, and variable inputs, v, such as labor and materials. This section describes our approach for estimating the output-to-capital elasticity ε_{tf}^{k} and the output-to-variable-input elasticity ε_{tf}^{v} from firm-level data on revenue (y), expenditures in variable inputs (v), and capital (k). Following Olley and Pakes (1996) and Ackerberg, Caves and Frazer (2015), we make the following assumptions:[41]

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 y_{tf}^{R} is given by a revenue production function of the form

$$\ln y_{tf}^{R} = z_{tf}^{R} + \varepsilon_{tc(f)}^{Rv} \cdot \ln v_{tf} + \varepsilon_{tc(f)}^{Rk} \cdot \ln k_{tf} + \epsilon_{tf},$$

where $c(f)$ denotes groups of firms with the same degree of automation and facing a common process for their revenue productivity, which only differs in their revenue productivity, z_{tf}^{R}, and an ex-post shock ϵ_{tf} that is orthogonal to k_{tf} and v_{tf}.

A3 Unobserved productivity z_{tf}^{R} evolves according to a Markov process of the form

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

[41] An alternative approach to estimating markups assumes constant returns to scale (as we do) directly measures 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. This strikes us as restrictive when thinking about firms undergoing a costly automation process. Instead, the approach described below makes no assumptions about the marginal product of capital across firms, or the importance of adjustment costs.

where ζ_{tf} is orthogonal to k_{tf} and v_{ft-1}, and the function g is common to all firms in the same group $c(f)$.

A4 True revenue, $\ln y_{tf}^{R*} = \ln y_{tf} - \epsilon_{tf}$ can be expressed as

$$\ln y_{tf}^{R*} = h(\ln x_{ft}, \ln k_{tf}, \ln v_{tf}),$$

where $\ln x_{tf} = \ln k_{t+1,f} - \ln k_{tf}$ denotes the investment rate of a firm and the function h is common to all firms in the same group $c(f)$.

A5 The gross output production function exhibits constant returns to scale in capital and variable inputs, which implies that output elasticities are given by

$$\varepsilon_{tf}^{v} = \varepsilon_{tc(f)}^{Rv} \Big/ \left(\varepsilon_{tc(f)}^{Rv} + \varepsilon_{tc(f)}^{Rk}\right) \qquad \varepsilon_{tf}^{k} = \varepsilon_{tc(f)}^{Rk} \Big/ \left(\varepsilon_{tc(f)}^{Rk} + \varepsilon_{tc(f)}^{Rk}\right). \tag{6}$$

Assumptions A1–A4 are standard in the literature. Assumption A4 justifies the use of the investment rate as a proxy variable. Economically, this assumption requires that all firms in a given group share the same investment policy function $k_{t+1,f} = \pi(k_{tf}, z_{tf}^{R})$, and that this common policy function is invertible. Under these assumptions, and given a grouping of firms $c(f)$, we can estimate revenue elasticities following the usual approach from Ackerberg, Caves and Frazer (2015), which uses the investment rate as a proxy variable to obtain true revenue and then estimates revenue elasticities by exploiting the orthogonality of ζ_{tf} to k_{tf} and $v_{t-1,f}$. Assumption A5 is added to deal with the fact that we don't observe prices, and so the usual estimation procedure yields revenue elasticities, not the quantity elasticities that are relevant for computing markups (Bond et al., 2020).[42] Under Assumption A5 we can recover output elasticities from revenue elasticities using (6).

We implement this approach using data from Compustat. Appendix F describes our sample selection and the details of our estimation approach. In our baseline approach, 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 firms operate different technologies and face a different demand curve, we group firms by quintiles of sales in each industry.[43] Thus, our estimation provides output

[42] Suppose that revenue is given by $y^R = p(q) \cdot q$, where $p(q)$ is the inverse demand curve. Quantity elasticities and revenue elasticities are then linked according to $\varepsilon^{Rv} = (p'(q) \cdot q/p(q) + 1) \cdot \varepsilon^{v}$ and $\varepsilon^{Rk} = (p'(q) \cdot q/p(q) + 1) \cdot \varepsilon^{k}$, where $1/\mu = (p'(q) \cdot q/p(q) + 1)$. Assuming constant returns to scale implies that $\varepsilon^{v} = \varepsilon^{Rv}/(\varepsilon^{Rv} + \varepsilon^{Rk})$, as wanted.

[43] 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_{tf}^{R} as true revenue);

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 the same output elasticities.[44]

3.2 Output elasticities and markups among Compustat firms

Figure 4 plots the estimated output-to-capital elasticities by size bin and time period. The left panel reports averaged elasticities across manufacturing industries, while the right panel reports averaged elasticities across non-manufacturing industries. In the 60s and 70s, firms had similar output-to-capital elasticities ranging from 0.08 to 0.12 in both sectors. For the following decades, we estimate a pronounced increase in output-to-capital elasticities among the firms in the top quintiles of the sales distribution. For the largest firms in manufacturing, the output-to-capital elasticity increases by 0.2 points, going from 0.11 to 0.31. This counter-clockwise rotation is precisely what our model predicts. In fact, our model generates an increase in the output-to-capital elasticity for large firms in manufacturing of 0.22 points. The fact that larger firms have become more capital intensive over time also aligns with the motivating evidence discussed in the introduction and which pointed to the uneven adoption of automation technologies by large firms.

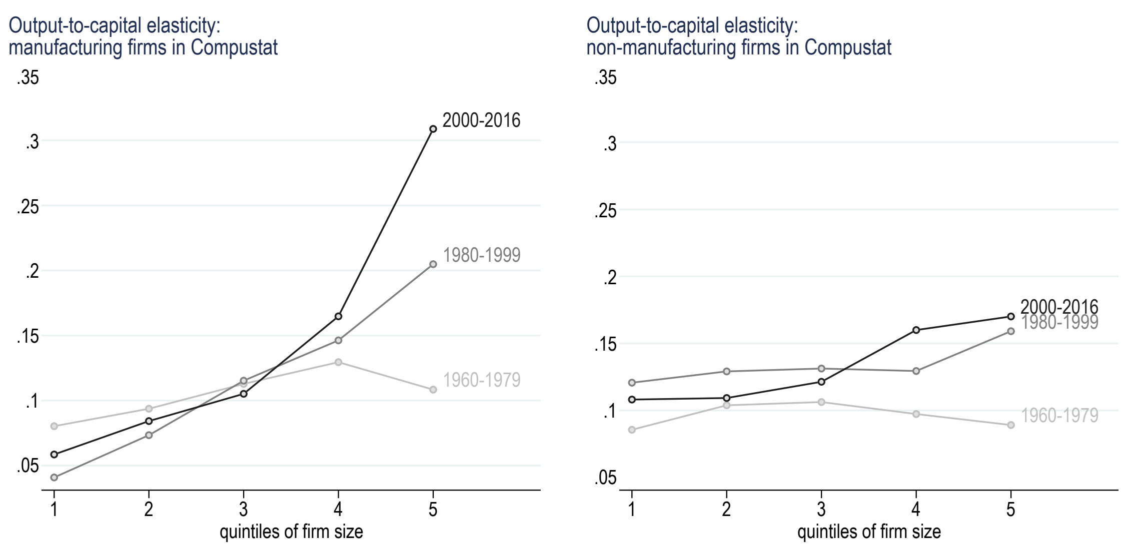

FIGURE 4: OUTPUT-TO-CAPITAL ELASTICITIES FOR COMPUSTAT FIRMS. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms.

or iii. assuming that z_{tf}^R 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.

[44]A byproduct of this estimation procedure are series for revenue TFP, z_{tf}^R. The estimated persistence of revenue TFP is 0.95 for manufacturing and 0.86 for retail, wholesale, utilities and transportation. These justifies the values of ρ_z used in our calibration approach.

Outside of manufacturing, we estimate a less pronounced increase in output-to-capital elasticities for large firms of 0.08 points (from 0.09 to 0.17). This is in line with the fact that, in our model, lower capital prices play a less prominent role in retail, wholesale, utilities and transportation. Indeed, our model generates an increase in output-to-capital elasticities for large firms in retail of 0.08 points.

We now turn to markups. Following Hall (1988), we estimate markups for a firm f at time t as $\mu_{tf} = \frac{\varepsilon^v_{tf}}{s^v_{tf}}$, where s^v_{tf} denotes the share of variable input expenditures in revenue. Figure 5 plots the implied time series for markups. We provide our estimates for 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. Our estimates for markups suggest that they have been quite stable over time at around 1.2.

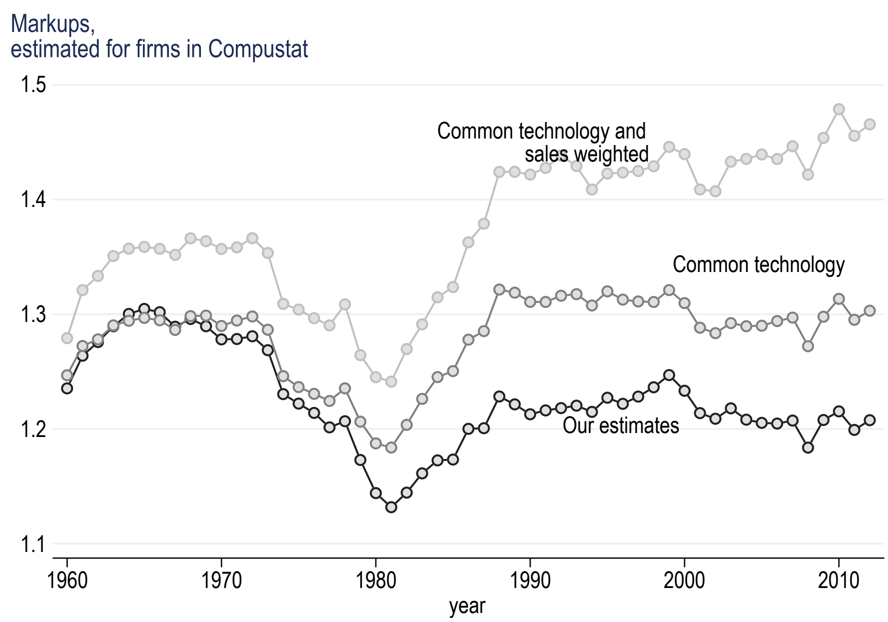

FIGURE 5: EVOLUTION OF MARKUPS. The figure presents the aggregate markup for firms in Compustat. Our estimates are obtained as as a sales-weighted *harmonic* mean of firm-level markups. The figure also reports the aggregate markup that would result under the assumption of common output elasticities across firms in the same industry, and a version of these estimates that aggregates firms' markups using a sales-weighted arithmetic mean.

For comparison, we provide the aggregate markup that would result if we assumed that all firms in the same industry operated technologies with the same capital intensity independently of their size class. 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 estimates in Edmond, Midrigan and Xu (2018). Finally, we provide estimates for 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 coincide with those reported by De Loecker, Eeckhout and Unger (2020). As explained above, the arithmetic mean of markups is irrelevant for understanding the contribution of markups to the decline in the labor share.

We now investigate the contribution of changes in firm markups and the reallocation across firms with different markups to the labor share decline in manufacturing and outside of manufacturing for 1980–2012. As discussed above, the labor share in an industry at time t can be written as ε_t^ℓ/μ_t, where ε_t^ℓ denotes the share of labor in industry costs and μ_t is a sales-weighted harmonic mean of firms' markups in the industry. Figure 6 plots the cumulative percent change in the inverse markup $1/\mu_t$ in red averaged across 3-digit manufacturing industries in the left panel and 3-digit non-manufacturing industries in the right panel. Furthermore, following the exercises in Section 2.3, we decompose this change into within-firm changes and the reallocation component—the contribution of reallocation towards firms with higher markups over time to the decline in the labor share.[45]

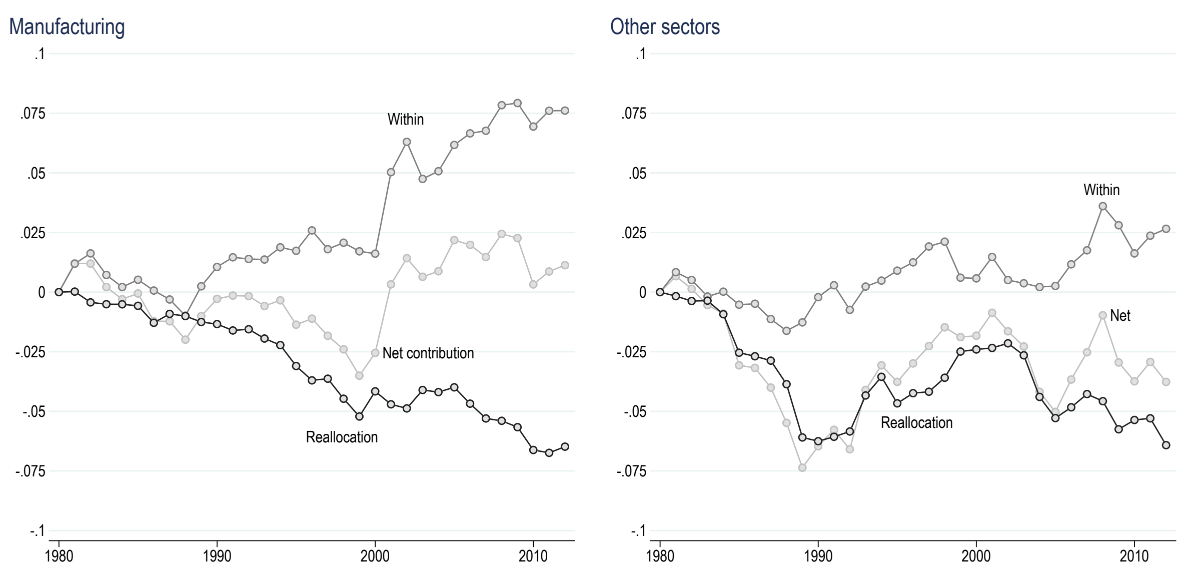

FIGURE 6: DECOMPOSITION OF THE CONTRIBUTION OF WITHIN-FIRM CHANGES IN MARKUPS AND BETWEEN-FIRM REALLOCATION TO (PERCENT) CHANGES IN THE LABOR SHARE. See the main text for details on this decomposition. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other economic sectors.

In line with our quantitative results, the reallocation to high-markup firms played a

[45]In 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 report the cumulative contribution of within firm changes and the reallocation component over time.

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. For comparison, our model generates a 3.6% decline in the labor share due to the reallocation towards large and high markup firms.[46] More importantly, within-firm changes in markups fully offset the reallocation component in manufacturing. This is in line with the fact that in our model the within firm changes in markups and the reallocation component roughly cancel each other out (a consequence of the close to Pareto productivity distribution in manufacturing). In sum, our estimates from Compustat support the conclusion that markups played a small role in driving the decline of the manufacturing labor share.

Outside of manufacturing, the reallocation component reduced the labor share by 6% since 1980, which accounts for half of the observed decline in retail. This is in line with our quantitative model, where we estimate a 6.2% reduction in the retail labor share due to the reallocation of activity towards high markup firms, most of it in response to rising competition. Moreover, in these sectors, the within component has been much weaker and close to zero. As a whole, our estimates for Compustat firms support the idea that, outside of manufacturing, rising competition might have reduced the labor share via reallocation towards high markup firms without bringing a similar offsetting reduction in firm markups.

4 Concluding remarks

The adoption of modern automation technologies concentrates at large firms. This paper shows that, once we account for this heterogeneity in technology adoption, one can explain the dynamics of the labor share decline across firms both qualitatively and quantitatively. 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. The model 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 US manufacturing sector. In particular, we find that the model reproduces the striking fact that while the manufacturing labor share declined drastically, the labor share of a typical firm increased or remained unchanged. The model also explains a range of other related observations, ranging from specific

[46] Our model only accounts for differences in markups driven by firm size. In principle, increased competition might also generate a reallocation of economic activity towards firms that are not necessarily large but have large markups for reasons that are not in our model, explaining the gap between model and data.

labor share decompositions proposed in the literature to the observed increases in concentration and productivity dispersion.

2. Second, we extended our model to an environment with endogenous markups, which allows us to account for the effects of rising competition and the ensuing reallocation to 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. The rise in competition is more important in retail and other sectors, where it accounts for up to 60% of the falling labor share.

3. Third, we contrasted the predictions of our model with empirical estimates of output elasticities and markups for US firms. Using standard techniques to estimate production functions but allowing for technology to vary across firms of different size, we estimate that the output-to-capital elasticity of large firms has increased over time, especially in manufacturing, which is exactly what our model predicts. By looking at empirical estimates of markups, we confirm that 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 (see Acemoglu et al., 2021). Moreover, our model points to the importance of developing estimators for markups and production functions 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 of markups and demand. Although the non-CES demand systems used here are gaining traction in macroeconomics, they restrict 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. 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 in such environments.

References

Acemoglu, D., G. Anderson, D. Beede, C. Buffington, E. Dinlersoz, L. Foster, N. Goldschlag, J. Haltiwanger, Z. K. P. Restrepo, and N. Zolas (2021): "Technology and the Size and Compositionof Workforce in U.S. Firms: First Evidence from the 2019 Annual Business Survey," Discussion paper, US Census.

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.

——— (2019): "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 33(2), 3–30.

——— (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.

Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2019): "A Theory of Falling Growth and Rising Rents," Working Paper 26448, National Bureau of Economic Research.

Akcigit, U., and S. T. Ates (2019): "What Happened to U.S. Business Dynamism?," Working Paper 25756, National Bureau of Economic Research.

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.

Caselli, F., and I. Coleman, Wilbur John (2006): "The World Technology Frontier," *American Economic Review*, 96(3), 499–522.

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.

De Ridder, M. (2020): "Market Power and Innovation in the Intangible Economy," Discussion paper, London School of Economics.

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.

Eldridge, L. P., C. Garner, T. F. Howells, B. C. Moyer, M. Russell, J. D. Samuels, E. H. Strassner, and D. B. Wasshausen (2020): ""Toward a BEA-BLS integrated industry-level production account for 1947–2016," in *Measuring Economic Growth and Productivity*, ed. by B. M. Fraumeni, chap. 11, pp. 221–249. Academic Press.

Elsby, M. W., B. Hobijn, and A. Şahin (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.

Greenwood, J., Z. Hercowitz, and P. Krusell (1997): "Long-Run Implications of Investment-Specific Technological Change," *The American Economic Review*, 87(3), 342–362.

Gutierrez, G., and S. Piton (2020): "Revisiting the Global Decline of the (Non-Housing) Labor Share," *American Economic Review: Insights*, (forthcoming).

Hall, R. (1988): "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, 96(5), 921–47.

Hall, R. E., and D. W. Jorgenson (1967): "Tax Policy and Investment Behavior," *The American Economic Review*, 57(3), 391–414.

Hopenhayn, H., J. Neira, and R. Singhania (2018): "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," Working Paper 25382, National Bureau of Economic Research.

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 Func-

tion 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. Santaeulalia-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.
Mariscal, A. (2020): "Firm Organization and Information Technology: Micro and Macro Implications," Discussion paper, Unpublished manuscript, U. Rovira i Virgili.
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.
Olmstead-Rumsey, J. (2019): "Market Concentration and the Productivity Slowdown," Discussion paper, Unpublished manuscript, London School of Economics.
Perla, J., C. Tonetti, and M. E. Waugh (2021): "Equilibrium Technology Diffusion, Trade, and Growth," *American Economic Review*, 111(1), 73–128.
Philippon, T. (2019): *The Great Reversal: How America Gave Up on Free Markets*. Harvard University Press.
Piketty, T. (2014): *Capital in the Twenty-First Century*. Harvard University Press, Cambridge.
Rossi-Hansberg, E., P.-D. Sarte, and N. Trachter (2021): "Diverging trends in national and local concentration," *NBER Macroeconomics Annual*, 35(1), 115–150.
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.
Zeira, J. (1998): "Workers, Machines, and Economic Growth," *The Quarterly Journal of Economics*, 113(4), 1091–1117.

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

Joachim Hubmer and Pascual Restrepo

December 18, 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 1–2 in the main text. In addition, we provide an additional lemma characterizing the stationary equilibrium of the economy.

Proof of Lemma 1. We first show that $\tilde{\alpha}_t(\alpha_{tf}, z)$ is weakly increasing in z. We have

$$\tilde{\alpha}_t(\alpha_{tf}, z) = \arg\max_{\alpha\in[\alpha_{tf},1]} -c_a \cdot y_t \cdot (\alpha - \alpha_{tf}) + \frac{1}{1+r}\mathbb{E}[V_{t+1,f}|z_{tf} = z, \alpha_{t+1,f} = \alpha].$$

It is therefore sufficient to show that $\mathbb{E}[V_{t+1,f}|z_{tf} = z, \alpha_{t+1,f} = \alpha]$ has increasing differences in (α, z). Ignore firm subscripts for simplicity, let $\Omega_{t+1}(\alpha, z) = \partial_\alpha \mathbb{E}[V_{t+1,f}|z_{tf} = z, \alpha_{t+1,f} = \alpha]$, and let $\pi_{t+1}(\alpha)$ denote the profits of a firm with automation level α and unitary productivity. The envelope theorem implies that

$$\Omega_t(\alpha, z) = \pi_t'(\alpha) \cdot \mathbb{E}[z'^{\sigma-1}|z] + \mathbb{E}\left[P_t(z') \cdot \min\left\{c_a \cdot y_t, \frac{1}{1+r}\Omega_{t+1}(\alpha, z')\right\}\Big|z\right], \tag{7}$$

where $P_t(z')$ denotes the probability of survival given z', and the minimum operator accounts for the fact that the restriction $\alpha_{t+1,f} \geq \alpha_{t,f}$ will bind in some states.

For every (t, α), define the following sequence:

$$\begin{aligned}
\Omega_t^{(1)}(\alpha, z) =& \pi_t'(\alpha) \cdot \mathbb{E}[z'^{\sigma-1}|z] \\
\Omega_t^{(n+1)}(\alpha, z) =& \pi_t'(\alpha) \cdot \mathbb{E}[z'^{\sigma-1}|z] + \mathbb{E}\left[P_t(z') \cdot \min\left\{c_a \cdot y_t, \frac{1}{1+r}\Omega_{t+1}^{(n)}(\alpha, z')\right\}\Big|z\right].
\end{aligned}$$

We prove by mathematical induction in n that, for all (t, α), $\Omega_t^{(n)}(\alpha, z)$ is weakly increasing in z. The base case for $n = 1$ follows from the fact that $\mathbb{E}[z'^{\sigma-1}|z]$ increases in z and $\pi_t'(\alpha) \geq 0$ (since firms can always choose to produce automated tasks with labor, and so a larger α weakly reduces their cost). For the inductive step, suppose that $\Omega_t^{(n)}(\alpha, z)$ is weakly increasing in z for all (t,α) with $n \le N$. We have

$$\Omega_t^{(N+1)}(\alpha,z) = \pi_t'(\alpha)\cdot\mathbb{E}[z'^{\sigma-1}|z] + \mathbb{E}\left[P_t(z')\cdot\min\left\{c_a\cdot y_t, \frac{1}{1+r}\Omega_{t+1}^{(N)}(\alpha,z')\right\}\Big|z\right].$$

As before, we have that $\pi_t'(\alpha)\cdot\mathbb{E}[z'^{\sigma-1}|z]$ is weakly increasing in z. Moreover, $P_t(z')\cdot\min\{c_a\cdot y_t, (1/(1+r))\cdot\Omega_{t+1}^{(N)}(\alpha,z')\}$ is (weakly) increasing in z' (due to the inductive hypothesis), and so the term $\mathbb{E}\left[P_t(z')\cdot\min\{c_a\cdot y_t,(1/(1+r))\cdot\Omega_{t+1}^{(N)}(\alpha,z')\}|z\right]$ also (weakly) increases in z, which completes the inductive step.

Because the set of weakly increasing functions is closed, $\Omega_t(\alpha,z) = \lim_{n\to\infty}\Omega_t^{(n)}(\alpha,z)$ is also weakly increasing in z. It follows that $\mathbb{E}[V_{t+1,f}|z_{tf} = z, \alpha_{t+1,f} = \alpha]$ has increasing differences in (α,z) as wanted.

Note that optimal automation decisions are guided by $\Omega_{t+1}(\alpha,z)$, which gives the marginal benefit to the firm of automating tasks up to $\alpha_{t+1,f} = \alpha$. Suppose that $\alpha_{tf} < \alpha^*_{t+1}$, and take any $\alpha \in [\alpha_{tf}, \alpha^*_{t+1})$, so that $\pi_{t+1}(\alpha)' > 0$. We assumed that, for any increasing and unbounded function f, $\mathbb{E}[f(z_{t+1,f})|z_{tf}]$ converges to infinity when $z_{tf}\to\infty$. This assumption implies that the right-hand side of equation (7) converges to infinity as $z\to\infty$. Thus, as $z\to\infty$, the optimal policy involves $\alpha = \alpha^*_{t+1}$, which is the only way to ensure that $\pi_{t+1}(\alpha)' = 0$ and $\Omega_{t+1}(\alpha,z) = 0$. Likewise, we assumed that, for any increasing function f, $\mathbb{E}[f(z_{t+1,f})|z_{tf}]$ converges to $f(0)$ when $z_{tf}\to 0$. Thus, the right-hand side of equation (7) converges to zero as $z\to 0$, which implies that $\Omega_{t+1}(\alpha,z) = 0$ for all α. In this case, the optimal policy is to keep $\alpha = \alpha_{tf}$ unchanged. ■

The following lemma will be used in our next results. Before turning to the lemma, we define a series of objects. First, given a constant path for capital prices, $q_t(x) = q(x)$, denote by $\alpha^*(w;q)$ the optimal level of automation for firms that face no costs of automation and face a wage w. As in the main text, this level is defined implicitly as

$$\frac{\psi^\ell(\alpha^*(w;q))}{q(\alpha^*(w;q))\cdot\psi^k(\alpha^*(w;q))} = w.$$

Finally, we let $w^*(q)$ denote the stationary equilibrium wage in a standard firm-dynamics model with no automation decisions, but with firms costs given by $c_{tf} = \frac{1}{z_{tf}}\cdot c(w;q)$, where the common cost function satisfies

$$c(w;q) = \left(\Psi^k(\alpha^*(w;q)) + \Psi^\ell(\alpha^*(w;q))\cdot w^{1-\eta}\right)^{\frac{1}{1-\eta}}.$$

The existence and uniqueness of this stationary equilibrium wage is given in Hopenhayn (1992).

Lemma 2 *Given a constant level of capital prices $q_t(x) = q(x)$, the economy admits a unique stationary equilibrium wage $w^*(q)$. Moreover, in any stationary equilibrium, $\alpha_{tf} \geq \alpha^*(q, w^*(q))$ for all firms, which implies that the economy behaves as if all firms had a unique level of automation $\alpha_{tf} = \alpha^*(q', w^*(q'))$.*

Proof. Consider a steady state with wage w. We first show that $\lim_{t\to\infty} \alpha_{tf} \geq \alpha^*(w;q)$. Consider the path for $\bar{\alpha}_t$. Because this is bounded from below, it must eventually lie in an ergodic set with infimum $\bar{\alpha}_\infty$. Suppose by way of contradiction that $\bar{\alpha}_\infty < \alpha^*(w;q)$. For large t, all entrants start with $\alpha_{tf} \geq \bar{\alpha}_\infty$, and they can only increase their α_{tf} over time. In fact, for any $\bar{\alpha}_\infty < \alpha^*(w;q)$, there will be a positive mass of entrants that will draw large realizations of z_{tf} through their lives, and will increase their α_{tf} strictly above $\bar{\alpha}_\infty$. This gives a contradiction, since the average α_{tf} would then exceed $\bar{\alpha}_\infty$ for all large t. This contradiction implies that $\bar{\alpha}_\infty \geq \alpha^*(w;q)$, as claimed.

Because $\bar{\alpha}_\infty \geq \alpha^*(w;q)$, all firms start with $\alpha_{tf} \geq \alpha^*(w;q)$ and retain this level of automation, producing only the tasks in $[0, \alpha^*(w;q)]$ with capital. The economy thus converges to a standard firm-dynamics model where firms costs are given by $c_{tf} = \frac{1}{z_{tf}} \cdot c(w;q')$. The unique steady-state equilibrium of this model then features a wage $w^*(q)$ and an automation level $\alpha_{tf} \geq \alpha^*(q, w^*(q))$ as claimed, but only tasks in $[0, \alpha^*(q, w^*(q))]$ are produced with capital. ■

Note: Lemma 2 justifies our focus on steady states where all firms operate a technology $\alpha^*(w^*(q);q)$ and wages are given by $w^*(q)$. Propositions 1 and 2 explore how factor shares vary across these steady states in response to different changes in capital prices.

Proof of Proposition 1. Let's write $q(x) = q \cdot q_0(x)$ and consider a permanent increase in q by $d\ln q$. 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.

Denote the mass of firms with productivity $z_{tf} = z$ by $m_z \geq 0$. We first show that, as q changes, the unit cost of production $\tilde{c}(w, q, \alpha^*)$ and the distribution of productivity among incumbents m_z remains unchanged across steady states. We prove this by showing that such an outcome satisfies the required steady-state equilibrium conditions. In particular, suppose that $\tilde{c}$ remains unchanged. Firm profits are then given by

$$\pi_{tf} = y \cdot \mu^{-\sigma} \cdot (\mu - 1) \cdot \tilde{c}(w, q, \alpha^*)^{1-\sigma} \cdot z^{\sigma-1},$$

which shows that firm profits are proportional to y, and so for a given z profits scale with aggregate output. Because fixed costs are also scaled by y, the value function of firms 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 m_z remains unchanged. The ideal-price index condition then implies that

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

which pins down the constant level for $\tilde{c}(w, q, \alpha^*)$.

We now use the fact that the unit cost function $\tilde{c}(w, q, \alpha^*)$ must be constant across steady-states to characterize the equilibrium response of wages. An application of Shephard's lemma implies that

$$d \ln \tilde{c}(w, q, \alpha^*) = \varepsilon^{\ell} \cdot d \ln w - \varepsilon^{k} \cdot d \ln q,$$

where, in addition, the envelope theorem ensures that the effect of changes in α^* on $\tilde{c}(w, q, \alpha^*)$ are second order and can be ignored. Because $d \ln \tilde{c}(w, q, \alpha^*) = 0$, 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 (or equivalently, output elasticities). In steady state, 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 (2) 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^*, we can compute the change in the cost share of labor as

$$
\begin{aligned}
d\ln\varepsilon^{\ell} &= \varepsilon^{k}\cdot d\ln\frac{\varepsilon^{\ell}}{\varepsilon^{k}} \\
&= \varepsilon^{k}\cdot(1-\eta)\cdot(d\ln q + d\ln w) + \varepsilon^{k}\cdot\frac{\partial\ln\Psi^{\ell}(\alpha^*)/\Psi^{k}(\alpha^*)}{\partial\ln\alpha}\cdot d\ln\alpha^* \\
&= \varepsilon^{k}\cdot(1-\eta)\cdot(d\ln q + d\ln w) + \varepsilon^{k}\cdot(\eta-\eta^*)\cdot(d\ln q + d\ln w) \\
&= \varepsilon^{k}\cdot(1-\eta^*)\cdot(d\ln q + d\ln w),
\end{aligned}
$$

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

$$
d\ln\varepsilon^{\ell} = \frac{1-\varepsilon^{\ell}}{\varepsilon^{\ell}}\cdot(1-\eta^*)\cdot d\ln q.
$$

Along the transition, firms will differ in the extent to which they will automate their tasks. Let $d\ln\alpha_{tf}$ denote the additional tasks automated by firm f at time t in response to the permanent increase in capital prices. We have that

$$
\begin{aligned}
d\ln\varepsilon^{\ell}_{tf} &= \varepsilon^{k}\cdot d\ln\frac{\varepsilon^{\ell}}{\varepsilon^{k}} \\
&= \varepsilon^{k}\cdot(1-\eta)\cdot(d\ln q + d\ln w_t) + \varepsilon^{k}\cdot\frac{\partial\ln\Psi^{\ell}(\alpha^*)/\Psi^{k}(\alpha^*)}{\partial\ln\alpha}\cdot d\ln\alpha_{tf} \\
&= \varepsilon^{k}\cdot(1-\eta)\cdot(d\ln q + d\ln w_t) + \varepsilon^{k}\cdot\frac{\partial\ln\Psi^{\ell}(\alpha^*)/\Psi^{k}(\alpha^*)}{\partial\ln\alpha}\cdot d\ln\alpha^*\cdot\frac{d\ln\alpha_{tf}}{d\ln\alpha^*} \\
&= \varepsilon^{k}\cdot(1-\eta)\cdot(d\ln q + d\ln w_t) + \varepsilon^{k}\cdot(\eta-\eta^*)\cdot(d\ln q + d\ln w_t)\cdot\frac{d\ln\alpha_{tf}}{d\ln\alpha^*} \\
&= \varepsilon^{k}\cdot\left(1-\eta+(\eta^*-\eta)\frac{d\ln\alpha_{tf}}{d\ln\alpha^*}\right)\cdot(d\ln q + d\ln w_t),
\end{aligned}
$$

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

$$
d\ln\varepsilon^{\ell}_{tf} = \frac{\varepsilon^{k}}{\varepsilon^{\ell}}\cdot\left(1-\eta+(\eta^*-\eta)\frac{d\ln\alpha_{tf}}{d\ln\alpha^*}\right)\cdot d\ln q.
$$

■

Proof of Proposition 2. Let $q(x) = q\cdot q_0(x)$ 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 1 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. In steady state, 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 (2) 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^*, we can compute the change in the cost share of labor as

$$\begin{aligned}
d\ln\varepsilon^{\ell} &= \varepsilon^{k} \cdot d\ln\frac{\varepsilon^{\ell}}{\varepsilon^{k}} \\
&= \varepsilon^{k} \cdot (1-\eta) \cdot d\ln w + \varepsilon^{k} \cdot \frac{\partial \ln \Psi^{\ell}(\alpha^*)/\Psi^{k}(\alpha^*)}{\partial \ln \alpha} \cdot d\ln\alpha^* \\
&= \varepsilon^{k} \cdot (1-\eta) \cdot d\ln w - \varepsilon^{k} \cdot (\eta^* - \eta) \cdot (d\ln q + d\ln w).
\end{aligned}$$

Along the transition, firms will differ in the extent to which they automate their tasks. Let $d\ln\alpha_{tf}$ denote the additional tasks automated by firm f at time t in response to the permanent increase in capital prices. We have that

$$\begin{aligned}
d\ln\varepsilon^{\ell}_{tf} &= \varepsilon^{k} \cdot d\ln\frac{\varepsilon^{\ell}}{\varepsilon^{k}} \\
&= \varepsilon^{k} \cdot (1-\eta) \cdot d\ln w_t + \varepsilon^{k} \cdot \frac{\partial \ln \Psi^{\ell}(\alpha^*)/\Psi^{k}(\alpha^*)}{\partial \ln \alpha} \cdot d\ln\alpha_{tf} \\
&= \varepsilon^{k} \cdot (1-\eta) \cdot d\ln w_t + \varepsilon^{k} \cdot \frac{\partial \ln \Psi^{\ell}(\alpha^*)/\Psi^{k}(\alpha^*)}{\partial \ln \alpha} \cdot d\ln\alpha^* \cdot \frac{d\ln\alpha_{tf}}{d\ln\alpha^*} \\
&= \varepsilon^{k} \cdot (1-\eta) \cdot d\ln w_t - \varepsilon^{k} \cdot (\eta^* - \eta) \cdot (d\ln q + d\ln w) \cdot \frac{d\ln\alpha_{tf}}{d\ln\alpha^*}.
\end{aligned}$$

■

A.2 The induced elasticity of substitution η^*

This section derives the parametrization of the productivity schedule for labor and capital that yields a constant induced elasticity of substitution between capital and labor of $\eta^* > 1$. For simplicity, we take $q(x) = q$ as in Proposition 1. Let $x_w = q \cdot w$ denote the wage relative to capital prices. Define $\alpha^*(x_w)$ implicitly as in the text by the solution to

$$x_w = \frac{\psi^\ell(\alpha^*(x_w))}{\psi^k(\alpha^*(x_w))}.$$

Define $h_k(x_w) = \Psi^k(\alpha^*(x_w)) \cdot q^{1-\eta}$ and $h_\ell(x_w) = \Psi^\ell(\alpha^*(x_w))$. Differentiating these expressions yields

$$h_k'(x_w) = \frac{\partial \alpha^*(x_w)}{\partial x_w} \cdot \psi^k(\alpha^*(x_w))^{\eta-1} \qquad h_\ell'(x_w) = -\frac{\partial \alpha^*(x_w)}{\partial x_w} \cdot \psi^\ell(\alpha^*(x_w))^{\eta-1},$$

which combined yield the differential equation

$$h_\ell'(x_w) = -x_w^{\eta-1} \cdot h_k'(x_w). \tag{8}$$

Recall that the long-run elasticity of substitution is defined implicitly by the identity

$$d \ln \frac{\varepsilon^\ell}{\varepsilon^k} \equiv (1-\eta^*) \cdot d \ln x_w.$$

In our model, the ratio of labor to capital in costs for a relative wage x_w can be computed as

$$\frac{\varepsilon^\ell}{\varepsilon^k} = \frac{h_\ell(x_w)}{h_k(x_w)} \cdot x_w^{1-\eta}.$$

It follows that the induced elasticity of substitution is constant and equal to η^* if and only if

$$\frac{h_\ell(x_w)}{h_k(x_w)} = \chi \cdot x_w^{\eta-\eta^*}. \tag{9}$$

This equation implies that $\eta^* > \eta$, since the left-hand side is decreasing in the wage. Rearranging this equation and taking derivatives yields

$$h_\ell'(x_w) = \chi \cdot x_w^{\eta-\eta^*} \cdot h_k'(x_w) - (\eta^* - \eta) \cdot \chi \cdot x_w^{\eta-\eta^*-1} \cdot h_k(x_w).$$

Combining this equation with (8) yields a differential equation for $h_k(x_w)$:

$$\frac{h_k'(x_w)}{h_k(x_w)} = (\eta^* - \eta) \cdot \frac{\chi \cdot x_w^{-\eta^*}}{\chi \cdot x_w^{1-\eta^*} + 1}. \tag{10}$$

This differential equation has two solutions, one for $\eta^* = 1$ and another one for $\eta^* > 1$. We will focus on the second one, and return to the first solution in Appendix C.3. Integrating both sides of equation 10 gives the unique solution for $h_k(x_w)$

$$h_k(x_w) = M \cdot \left(\chi \cdot x_w^{1-\eta^*} + 1\right)^{\frac{\eta-\eta^*}{\eta^*-1}}.$$

Using equation (9), we also obtain

$$h_\ell(x_w) = M \cdot \chi \cdot \left(\chi + x_w^{\eta^*-1}\right)^{\frac{\eta-\eta^*}{\eta^*-1}}$$

With the functions $h_k(x_w)$ and $h_\ell(x_w)$ at hand, we can now generate all possible parametrizations of the productivity schedules for capital and labor that induce a constant elasticity of substitution $\eta^* > 1$. In particular, for any increasing function $\alpha^*(x_w) = f(x_w)$ from the positive reals to $[0,1]$, we can define

$$\psi^k(x) = \left[\frac{h_k'(f^{-1}(x))}{f'(f^{-1}(x))}\right]^{\frac{1}{\eta-1}} \qquad \psi^\ell(x) = \left[-\frac{h_\ell'(f^{-1}(x))}{f'(f^{-1}(x))}\right]^{\frac{1}{\eta-1}}.$$

This parametrization yields $\Psi^k(x) = h_k(f^{-1}(x))$ and $\Psi^\ell(x) = h_\ell(f^{-1}(x))$, an optimal threshold rule given by $\alpha^*(x_w) = f(x_w)$, and an induced elasticity of substitution of η^*. The parametrization in the main text comes from taking the natural choice of $f(x_w) = h_k(x_w)$, which is an increasing function from the positive reals to $[0,1]$.

B Proofs for the model with variable markups

This section provides the proofs of the theoretical results in section 2.

B.1 Implications of Marshall's weak and strong second laws

We begin with a lemma that characterizes the implications of Marshall's second laws for prices, markups, and passthroughs. We consider a firm with a constant marginal cost c and denote its optimal price by $p^*(c)$. Likewise, we define markups by $\mu^*(c)$ and firm sales by $\omega^*(c)$.

Lemma 3 *Under Marshall's weak second law, firms with lower costs c charge lower prices $p^*(c)$ but higher markups $\mu^*(c)$. Moreover, under Marshall's strong second law, markups and prices, $\mu^*(c)$ and $p^*(c)$, are a log-convex function of costs, which implies lower passthroughs for more productive firms. Finally, sales $\omega^*(c)$ are a log-concave and decreasing function of costs.*

Proof of Lemma 3. 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

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

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

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

where

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

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 shares, we have that $\omega^*(c)$ can be written as

$$\omega^*(c) = h(p^*(c))/y,$$

where $h(x) = xD(x)$ is a log-concave and decreasing function of x (from Marshall's weak second law). Thus, $\omega^*(c)$ 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. ■

B.2 Proofs and derivations of results in the main text

Before turning to the proofs of the propositions in the text, we provide some preliminary derivations and the full formal definition of an equilibrium in the non-CES model.

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

$$\min_{y_{tf}} \int_f p_{tf} \cdot y_{tf} \cdot df \qquad \text{s.t:} \quad \int_f \lambda \cdot H\left(\frac{y_{tf}}{\lambda_t \cdot y_t}\right) \cdot df = 1.$$

Let $\rho_t \cdot y_t$ denote the Lagrange multiplier on the constraint. The first-order condition for the choice of y_{tf} is then

$$p_{tf} = \rho_t \cdot H'\left(\frac{y_{tf}}{\lambda_t \cdot y_t}\right) \quad \Rightarrow \quad y_{tf} = y_t \cdot \lambda_t \cdot D\left(\frac{p_{tf}}{\rho_t}\right).$$

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

$$1 = \int_f \lambda_t \cdot p_{tf} \cdot D\left(\frac{p_{tf}}{\rho_t}\right) \cdot df, \tag{11}$$

which is the ideal-price index condition for the non-CES model.

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

$$\int_f \lambda_t \cdot H\left(D\left(\frac{p_{tf}}{\rho_t}\right)\right) \cdot df = 1, \tag{12}$$

which pins down the competitors' price index ρ_t.

Denote by $p_{tf}(w)$ the price charged by a firm facing a wage w, by $c_{tf}(w)$ its cost, and by $\pi_{tf}(w)$ its profits. Given a path for investment productivities $q_t(x)$, market size, λ_t, and an initial distribution of firms $\{\alpha_{0f}, z_{0f}\}$, an equilibrium is given by a path for wages w_t, aggregate output y_t, the competitors' price index ρ_t, as well as a path for the distribution of firms $\{\alpha_{tf}, z_{tf}\}$, such that for all $t \geq 0$:

E1. The ideal-price index condition in equation (11) holds.

E2. The competitors' price index condition in equation (12) holds.

E3. The labor market clears

$$\int_f y_t \cdot \lambda_t \cdot D\left(\frac{p_{tf}}{\rho_t}\right) \cdot \frac{\partial c_{tf}(w_t)}{\partial w_t} \cdot df = \ell.$$

E4. Automation and exit decisions maximize the value function of incumbents

$$
\begin{aligned}
V_{tf} =& \pi_{tf}(w_t) + \\
& \int \max\left\{0, -c_o \cdot y_t + \max_{\alpha_{t+1,f}\in[\alpha_{t,f},1]} \left\{-c_a \cdot y_t \cdot (\alpha_{t+1,f} - \alpha_{t,f}) + \frac{1}{1+r}\mathbb{E}\left[V_{t+1,f}|z_{t,f}\right]\right\}\right\} dG(c_o).
\end{aligned}
$$

E5. Entry decisions maximize the value of entrants

$$
V^e_{tz} = \int \max\left\{0, -c_o \cdot y_t + \max_{\alpha_{t+1,f}\in[\bar{\alpha}_t,1]} \left\{-c_a \cdot y_t \cdot (\alpha_{t+1,f} - \bar{\alpha}_t) + \frac{1}{1+r}\mathbb{E}\left[V_{t+1,f}|z_{tf} = z\right]\right\}\right\} dG(c_o),
$$

where z denotes an entrant's productivity signal, and $\bar{\alpha}_t \equiv (\int_f \alpha_{tf} \cdot df)/(\int_f df)$.

E6. Starting from a distribution $\{\alpha_{0f}, z_{0f}\}$, the evolution of $\{\alpha_{tf}, z_{tf}\}$ is governed by the exogenous process for z, the endogenous process for α, and optimal entry and exit decisions.

Proof of Proposition 3. Let $\bar{c} = \widetilde{c}/\rho$, where recall that $\widetilde{c}$ is the constant marginal cost for a firm with unitary productivity. We can rewrite firms' pricing problem as

$$
\max_{\bar{p}} D(\bar{p}) \cdot (\bar{p} - \bar{c}/z),
$$

where $\bar{p} = p/\rho$ denotes the normalized firm price . Lemma 3 implies that firm prices are given by $p_z = \rho \cdot p^*(\bar{c}/z)$, markups by $\mu_z = \mu^*(\bar{c}/z)$, and sale shares by $\omega_z = \omega^*(\bar{c}/z)$.

The implicit definition of the competitors' price index implies

$$
\int_z \lambda \cdot H\left(D\left(p^*\left(\bar{c}/z\right)\right)\right) \cdot m_z \cdot dz = 1.
$$

Consider an increase in λ. Suppose by way of contradiction that $\bar{c}$ declines. This would increase firm profits, increasing entry and reducing exit. Note also that any effect of λ on aggregate output *holding* $\bar{c}$ constant will not affect entry or exit decisions. This is because, conditional on $\bar{c}$, value functions are linear in aggregate output y. As a result, m_z would increase and the price index condition would be violated. This contradiction then requires $\bar{c}$ to increase. As a result, the effect of an increase in λ on prices, markups, and sales shares can be summarized by the resulting increase in $\bar{c}$.

We now characterize the effects of an increase in $\bar{c}$. First, we have that for a given z, $\mu_z = \mu^*(\bar{c}/z)$ will be decreasing in $\bar{c}$, as wanted. Second, because the function $\mu^*(c)$ is log-convex, we have that, for $z > z'$,

$$
\ln \mu_z - \ln \mu_{z'} = \ln \mu^*\left(\bar{c}/z\right) - \ln \mu^*\left(\bar{c}/z'\right)
$$

is decreasing in $\bar{c}$. Third, because the function $\omega^*(c)$ is log-concave, we have that, for $z > z'$,

$$\ln \omega_z - \ln \omega_{z'} = \ln \omega^* (\bar{c}/z) - \ln \omega^* (\bar{c}/z')$$

is increasing in $\bar{c}$. ■

Proof of Proposition 4. As before, we investigate the implications of an increase in $\bar{c}$. Holding the distribution of productivities constant at $m_z = f(z)$, we can write the aggregate markup as

$$\frac{1}{\mu} = \int_z \frac{1}{\mu^* (\bar{c}/z)} \cdot \omega^* (\bar{c}/z) \cdot f(z) \cdot dz.$$

With the change of variable $x = \bar{c}/z$, 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

$$g(x, \bar{c}) = \omega^* (x) \cdot f(\bar{c}/x) \cdot \frac{\bar{c}}{x^2} \cdot dx.$$

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

$$\ln g(x, \bar{c}) = \ln \omega^* (x) + \ln f(\bar{c}/x) + \ln \bar{c} - 2 \ln x$$

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 $\frac{1}{\mu^*(x)}$ is increasing in x, the aggregate markup μ decreases in $\bar{c}$ as wanted.

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

$$\ln g(x, \bar{c}) = \ln \omega^* (x) + \ln f(\bar{c}/x) + \ln \bar{c} - 2 \ln x$$

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 $\frac{1}{\mu^*(x)}$ is increasing in x, the aggregate markup μ increases in $\bar{c}$ as wanted.

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

$$\ln g(x,\bar{c}) = \ln \omega^*(x) + \ln f(\bar{c}/x) + \ln \bar{c} - 2\ln x$$

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

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

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

B.3 Properties of the Klenow–Willis aggregator

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

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

where $\bar{y}_{tf} = y_{tf}/(\lambda_t \cdot y_t)$ 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 gives rise to the following (relative) demand function $D^{-1} = H'$:

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

where $\bar{p}_{tf} = p_{tf}/\rho$ is the normalized price of a variety. The price elasticity of demand is

$$-\frac{D'(\bar{p}_{tf}) \cdot \bar{p}_{tf}}{D(\bar{p}_{tf})} = \frac{\sigma}{1 - \nu \cdot \ln\left(\bar{p}_{tf} \cdot \frac{\sigma}{\sigma-1}\right)} = \sigma \cdot D(\bar{p}_{tf})^{-\frac{\nu}{\sigma}}, \tag{13}$$

which reduces to the constant σ if $\nu = 0$ (the benchmark case of a CES aggregator). In general, equation (13) shows that under this parametrization, the super-elasticity of demand is equal to the constant $-\frac{\nu}{\sigma}$, and that larger firms will face more inelastic demand curves.

To conlcude, we show that the Klenow-Willis aggregator satisfies Marshall's second laws. Equation (5) 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

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

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

C Additional quantitative results and robustness exercises

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. We also provide results for the model with non-CES demand in Section 2 with an alternative lower super-elasticity of demand.

C.1 Timing of automation decisions

In the main text, we assumed that firms invest in $\alpha_{t+1,f}$ in period t before the realization of their productivity in period $t+1$. We have experimented with other timing assumptions and found that our results are robust on this dimension. 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. 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). Because the option value of automation for the median firm is now lower, the automation fixed cost required to match the automation gradient decreases by 18%; in turn, firms with high z-innovations can automate immediately, which increases the importance of the cross-cross dynamics term slightly.

TABLE 7: Model robustness: Calibrations for alternative versions of the CES demand model in Section 1.2 (manufacturing)

		DATA	MODEL			
			BENCH-MARK	ALTERN. TIMING	LOWER η	LOWER η^*
		(1)	(2)	(3)	(4)	(5)
	I. Parameters					
η	Task substitution elasticity		0.50	0.50	0.25	0.50
γ	Comparative advantage		0.95	0.95	1.20	*
σ	Demand elasticity		7.67	7.67	7.67	7.67
$\underline{c}_o$	Scale operating cost ($\times 10^{-7}$)		4.0	4.0	4.0	4.0
ξ_o	Tail index operating cost		0.250	0.250	0.250	0.250
μ_e	Entrant productivity		0.905	0.905	0.905	0.905
σ_z	Std. dev. of $\ln z$ innovations		0.105	0.105	0.105	0.105
ρ_z	Productivity persistence		0.95	0.95	0.95	0.95
	II. Moments					
	Aggregate markup	1.15	1.15	1.15	1.15	1.15
	Top 4 firms' sales share	40.0%	40.0%	40.2%	40.0%	40.1%
	Entry (=exit) rate	0.062	0.063	0.063	0.063	0.063
	Size of exiters	0.490	0.490	0.497	0.490	0.491
	Size of entrants	0.600	0.600	0.589	0.600	0.598
	Task substitution elasticity		0.50	0.50	0.25	0.50
	Long-run K-L elasticity		1.45	1.45	1.45	1.00

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 after the realization of productivity 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 induced capital-labor elasticity η^*.

C.2 Short-run capital-labor elasticity

Column (4) in Table 7 describes an alternative calibration with $\eta = 0.25$. For this exercise, we hold the long-run elasticity constant, which requires increasing the comparative advantage parameter γ from $1.45 - 0.5 = 0.95$ to $1.45 - 0.25 = 1.2$. Comparing columns (2) and (4) in Table 8 reveals that the results largely coincide, implying that our findings are not sensitive to the value of η within a reasonable range. Since the comparative advantage of labor is flatter (γ is higher), the automation incentive is stronger even for the median firm; thus, we infer a 35% higher fixed cost of automation when targeting the same automation gradient. The dynamics of labor and market shares are mostly unaffected.

Table 8: Model robustness: Transitional dynamics for alternative versions of the CES demand model in Section 1.2 (manufacturing)

	Data	Model: Benchmark	Altern. timing	Lower η	$\eta^* < 1$
	(1)	(2)	(3)	(4)	(5)
I. Parameters and inferred aggregate shocks					
$d \ln q$	.	1.68	1.68	1.70	0.28*
c_a	.	0.354	0.290	0.480	0.135
II. Targeted moments, 1982–2012					
Δ aggregate labor share	−0.20	−0.20	−0.20	−0.20	−0.20
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1.71	1.70	1.71
III. Typical firm labor share (Kehrig and Vincent, 2020), 1982–2012					
Δ median labor share	0.030	−0.003	0.002	0.022	−0.013
Δ unweighted mean	−0.017	−0.023	−0.020	−0.002	−0.051
IV. Other moments, 1982–2012					
Δ log top 4 firms' sales share	0.140	0.071	0.057	0.074	0.072
Δ log top 20 firms' sales share	0.072	0.070	0.061	0.072	0.074
Δ log productivity dispersion	0.050	0.059	0.061	0.068	0.035
V. Melitz–Polanec decomposition from Autor et al. (2020)					
Δ aggregate labor share	−0.185	−0.201	−0.200	−0.203	−0.201
Δ unweighted incumbent mean	−0.002	0.006	0.006	0.031	−0.043
Exit	−0.055	−0.004	−0.004	−0.004	−0.003
Entry	0.059	0.006	0.005	0.007	0.004
Covariance term	−0.187	−0.209	−0.207	−0.236	−0.159
VI. Covariance decomposition from Kehrig and Vincent (2020)					
Market share dynamics	0.047	0	0	0	0
Labor share by size dynamics	−0.043	−0.111	−0.109	−0.096	−0.118
Cross-cross dynamics	−0.232	−0.095	−0.101	−0.112	−0.085

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 after the realization of productivity 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 induced capital-labor elasticity $\eta^* = 1$. The reported value of $d \ln q$ for this parameterization refers to the change in the average price of capital, computed using a Törnqvist index (see details in text).

C.3 Long-run capital-labor elasticity

In the main text, we studied the effects of a uniform decline in the price of capital. This shock requires an above one long-run capital-labor elasticity ($\eta^* > 1$) to match the decline in the labor share observed in the data. Following Proposition 2, here we show that even if $\eta^* \leq 1$, we can also generate a decline in the aggregate labor share and a constant labor share of the typical 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 $\eta^* \leq 1$.

In particular, we describe here a specification where the induced elasticity of substitution at the initial and the final steady state is exactly 1.

Parametrization of productivity schedule: We adopt the following functional forms for capital productivity, relative capital prices, and labor productivity:[47]

$$\psi^k(x) = x^{\frac{1-\gamma_k}{1-\eta}} \cdot (1-x)^{\frac{1+\gamma_k}{1-\eta}}, \qquad q_0(x) = 1, \qquad \psi^\ell(x) = x^{\frac{1+\gamma_\ell}{1-\eta}} \cdot (1-x)^{\frac{1-\gamma_\ell}{1-\eta}},$$

where $\gamma_k, \gamma_\ell > 0$. As before, the task space is the unit interval, and $\eta < 1$. Again, we consider an economy in steady state in 1982 with wage w_0. The implied share parameters are

$$\Psi^k(\alpha) = \frac{1}{\gamma_k}\left(\frac{\alpha}{1-\alpha}\right)^{\gamma_k}, \qquad \Psi^\ell(\alpha) = \frac{1}{\gamma_\ell}\left(\frac{\alpha}{1-\alpha}\right)^{-\gamma_\ell}.$$

Moreover, the optimal automation decision satisfies

$$\alpha_0^* = \left(\frac{w_0}{1+w_0}\right)^{\frac{1-\eta}{\gamma_\ell+\gamma_k}},$$

which implies that the ratio of labor to capital costs is given by

$$\frac{s_0^\ell}{s_0^k} = \frac{\Psi^\ell(\alpha_0^*)}{\Psi^k(\alpha_0^*)} \cdot w_0^{1-\eta} = \frac{\gamma_k}{\gamma_\ell}.$$

This shows that, once firms are allowed to adjust their tasks, the labor share in costs is constant and equal to $\gamma_k/(\gamma_k+\gamma_\ell)$. It follows that the induced elasticity of substitution in 1982 equals 1. In particular, an econometrician with data from 1982 and exploiting wage variation would conclude that firms operate a Cobb-Douglas production function. In what follows, we normalize $\gamma_k = 1$, so that γ_ℓ controls the labor share in 1982.

Falling capital prices at marginal tasks: In the spirit of Proposition 2, consider a gradual increase in $q(x)$ such that $q(x)$ goes from 1 to $\bar{q} \cdot \left(\frac{x}{1-x}\right)^{\frac{\gamma_q}{1-\eta}} > 1$ for $x \in (\alpha_0^*, 1]$. $\bar{q}$

[47] We derived this specification by solving (10) for $\eta^* = 1$. In this case, we can write (10) as

$$\frac{h_k'(w)}{h_k(w)} = (\eta^* - \eta) \cdot \frac{\chi}{\chi+1} \cdot \frac{1}{w}.$$

Integrating both sides we get $h_k(w) = M \cdot w^{(1-\eta)\cdot\frac{\chi}{\chi+1}}$ and $h_\ell(w) = M \cdot w^{(1-\eta)\cdot\frac{\chi}{\chi+1}}$. Taking the function $\alpha^*(x_w) = \left(\frac{x_w}{1+x_w}\right)^{\frac{1-\eta}{\gamma_\ell+\gamma_k}}$ yields the parametrization used here.

gradually increases over time, and the constant $\gamma_q \in (0,1)$ is chosen so that

$$\bar{q} \cdot \left(\frac{\alpha_0^*}{1-\alpha_0^*} \right)^{\frac{\gamma_q}{1-\eta}} = \left(\frac{1}{1-\gamma_q} \right)^{\frac{1}{1-\eta}},$$

which ensures that capital prices fall for all tasks above α_0^*, and that the increase is more pronounced for higher-indexed tasks. The restriction on $\bar{q}$ and γ_q ensures that the induced elasticity remains exactly at 1. The implied share parameters, given $\alpha \geq \alpha_0^*$, are

$$\Psi^k(\alpha) = \left(\frac{\alpha_0^*}{1-\alpha_0^*} \right) + \frac{\bar{q}^{\eta-1}}{1-\gamma_q} \left(\left(\frac{\alpha}{1-\alpha} \right)^{1-\gamma_q} - \left(\frac{\alpha_0^*}{1-\alpha_0^*} \right)^{1-\gamma_q} \right), \qquad \Psi^\ell(\alpha) = \frac{1}{\gamma_\ell} \left(\frac{\alpha}{1-\alpha} \right)^{-\gamma_\ell}.$$

Moreover, the optimal automation threshold in the new steady state is

$$\alpha_1^* = \left(\frac{\bar{q} \cdot w_1}{1 + \bar{q} \cdot w_1} \right)^{\frac{1-\eta}{\gamma_\ell + 1 - \gamma_q}},$$

which implies that the ratio of labor to capital costs in the new steady state is given by

$$\frac{s_1^\ell}{s_1^k} = \frac{\Psi^\ell(\alpha_1^*)}{\Psi^k(\alpha_1^*)} \cdot w_1^{1-\eta} = \frac{1-\gamma_q}{\gamma_\ell}.$$

As before, the induced elasticity of substitution is 1. The labor share in costs equals $(1-\gamma_q)/(1-\gamma_q+\gamma_\ell)$, which is lower than in 1982 (recall that $\gamma_q \in (0,1)$). Thus, even though lower capital prices at marginal tasks drive the decline in the labor share, an econometrician with data from the new steady state and exploiting wage variation would still conclude that firms operate a Cobb-Douglas production function.

Calibration: Column (5) in Table 7 describes the full steady state calibration for this version of the model. We now calibrate $\gamma_\ell = 0.30$ to match the labor share in 1982. Column (5) in Table 8 describes the transitional dynamics in response to a decrease in capital prices parametrized by q (and the associated γ_q), which we calibrate to match the aggregate labor share decline from 1982 to 2012 in manufacturing. We infer a value of $d \ln \bar{q} = 2.52$ and $\gamma_q = 0.67$, which jointly generate a small reduction in the average capital price of only 28 log points (computed using a Törnqvist index). This is because the price of capital at existing capital tasks, in the interval $[0, \alpha_0^*]$, is unchanged. For tasks re-allocated to capital, in the interval $\in (\alpha_0^*, \alpha_1^*]$, the price of capital falls by 220 to 508 log points. In addition, we infer a lower fixed cost of automation of $c_a = 0.14$ to match the adoption rates by firm size from Acemoglu et al. (2021).

Quantitative findings: The benchmark model with a uniform decline in capital prices and $\eta^* = 1.45$ (column (2) in Table 8) and the alternative one with falling capital prices at marginal tasks only and $\eta^* = 1$ described here (column (5) in Table 8) produce comparable results. While the labor share of the median firm is decreasing slightly (−1.3 pp) in the alternative specification, the firm-level labor and market share dynamics are comparable: 80% of the falling manufacturing labor share is accounted for by the covariance term (as opposed to 100% in data and benchmark model), and the model similarly accounts for half of the increase in the top 4 firms' sales share (and all of the increase in the top 20 firms' sales share). We conclude that, even when the aggregate production function is Cobb-Douglas, we can generate the observed decline in the aggregate labor share and the mostly flat labor share of the typical firm as a response to a reduction of capital prices at marginal tasks.

C.4 Super-elasticity of demand

In the main text, we calibrated a demand super-elasticity of $\frac{\nu}{\sigma} = 0.22$ by matching the ratio of the (unweighted) mean firm labor share to the aggregate sectoral labor share. Here, we report results for the main sectors of interest, manufacturing and retail, when instead using a lower super-elasticity of 0.16 as estimated by Edmond, Midrigan and Xu (2018). For this exercise, we re-calibrate the parameters in the initial steady states for both sectors. The main difference is that a lower value of the super-elasticity requires less convexity in the productivity distribution, since the mapping from productivity to firm sales is less log-concave. For manufacturing, we infer $n = 0.91$ (instead of $n = 0.74$ as in Table 3); for retail, we infer $n = 0.63$ (instead of $n = 0.47$ as in Table 4). Thus, the inferred z-distributions are closer to the log-linear Pareto distribution, which is the special case with $n = 1$.

Table 9 reports the main results over the transition (1982–2012) for both sectors. Relative to the results in the main text, the inferred rising competition shocks are somewhat larger: in manufacturing, we infer $d\ln\lambda = 0.32$ (instead of $d\ln\lambda = 0.29$ as in Table 5); in retail, we infer $d\ln\lambda = 0.48$ (instead of $d\ln\lambda = 0.30$ as in Table 6). Even though the inferred shocks are larger, the lower log-convexity of the z-distribution implies that the λ-shock generates a smaller increase in the aggregate markup, and correspondingly a smaller decrease in the aggregate labor share. As documented in Table 9, we infer a rise in the aggregate markup of 0.9% in manufacturing (as opposed to 1.3% in Table 5) and of 2.5% in retail (as opposed to 4.9% in Table 5). The other main results, in particular the contribution of falling capital prices to the labor share decline, are quantitatively very similar across parameterizations.

TABLE 9: Model robustness: Transitional dynamics in manufacturing and retail under a non-CES demand system with lower super-elasticity of $\frac{\nu}{\sigma} = 0.16$

		Model		
	Data	Benchmark	Only effects of $d \ln q$	Only effects of $d \ln \lambda$
	(1)	(2)	(3)	(4)
A. Manufacturing (1982–2012)				
I. Parameters and inferred aggregate shocks				
$d \ln q$	.	1.57	1.57	0
$d \ln \lambda$	.	0.09	0	0.09
c_a	.	0.32	0.32	0.32
II. Targeted moments, 1982–2012				
Δ aggregate labor share	−0.199	−0.200	−0.184	−0.002
Δ log top 4 firms' sales share	0.140	0.140	0.081	0.057
Δ Relative adoption (P99+ vs. P50-75)	1.71	1.71	1.61	5.63
III. Typical firm labor share and other moments, 1982–2012				
Δ median labor share	0.030	0.042	0.027	0.010
Δ unweighted mean labor share	−0.017	0.013	0.004	0.008
Δ log top 20 firms' sales share	0.072	0.129	0.091	0.036
Δ log productivity dispersion	0.050	0.069	0.059	0.000
IV. Markups, 1982–2012				
Δ log aggregate markup	.	0.009	0.009	0.000
Within firm change in markups	.	−0.019	−0.015	−0.011
Reallocation	.	0.028	0.023	0.010
B. Retail (1982–2012)				
I. Parameters and inferred aggregate shocks				
$d \ln q$	.	0.85	0.85	0
$d \ln \lambda$	.	0.48	0	0.48
c_a	.	0.36	0.36	0.36
II. Targeted moments, 1982–2012				
Δ aggregate labor share	−0.127	−0.127	−0.057	−0.040
Δ log sales concentration	0.546	0.541	0.031	0.498
Δ Relative adoption (P99+ vs. P50-75)	1.71	1.71	1.48	2.12
III. Typical firm labor share and other moments, 1982–2012				
Δ median labor share	.	0.048	0.001	0.032
Δ unweighted mean labor share	.	0.033	−0.004	0.027
Δ log productivity dispersion	.	0.021	0.007	0.002
IV. Markups, 1982–2012				
Δ log aggregate markup	.	0.025	0.004	0.020
Within firm change in markups	.	−0.023	−0.008	−0.020
Reallocation	.	0.048	0.013	0.040

Notes: Panels A and B report the equivalents of Tables 5 and 6 in the main text, when imposing instead a lower super-elasticity of $\frac{\nu}{\sigma} = 0.16$ (instead of $\frac{\nu}{\sigma} = 0.22$ as in the main text). The parameters of the respective economies are re-calibrated, both in the steady state to match all other targeted moments, as well as in regards to the inferred shocks $d \ln q$, $d \ln \lambda$ and the automation fixed cost c_a over the transition.

D Calibration of the non-CES demand model for other sectors

Table 10 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: Steady state calibration of the non-CES demand model: Wholesale, Utilities & Transportation

	Parameter		Moment	Data	Model
	I. Wholesale: steady state parameters and moments (1982)				
ν/σ	Demand super-elasticity	0.22	Imputed from manufacturing		
σ	Demand elasticity	9.4	Aggregate markup	1.15	1.15
ζ	Weibull scale	0.071	Top 20 firms' sales share	42.9%	42.9%
n	Weibull shape	0.75	Top 4 firms' sales share	22.3%	22.3%
$\underline{c}_o$	Scale operating cost	$3.2 \cdot 10^{-7}$	Entry (=exit) rate	0.062	0.062
ξ_o	Tail index operating cost	0.235	Size of exiters	0.490	0.493
μ_e	Entrant productivity	0.889	Size of entrants	0.600	0.601
ρ_z	Productivity persistence	0.86	Revenue TFP persistence among wholesale firms		
	II. Utilities & Transportation: steady state parameters and moments (1992)				
ν/σ	Demand super-elasticity	0.22	Imputed from manufacturing		
σ	Demand elasticity	10.7	Aggregate markup	1.15	1.15
ζ	Weibull scale	0.066	Top 20 firms' sales share	59.1%	58.0%
n	Weibull shape	0.74	Top 4 firms' sales share	30.4%	31.3%
$\underline{c}_o$	Scale operating cost	$9.0 \cdot 10^{-8}$	Entry (=exit) rate	0.062	0.063
ξ_o	Tail index operating cost	0.212	Size of exiters	0.490	0.489
μ_e	Entrant productivity	0.891	Size of entrants	0.600	0.600
ρ_z	Productivity persistence	0.86	Revenue TFP persistence among ut. & transp. firms		

Notes: 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 remaining data moments follow the model with CES demand, see Table 1. Fixing productivity persistence and the demand super-elasticity, in each of the two sectors the remaining six parameters are jointly calibrated to match the six corresponding moments.

Table 11 shows the model-based decomposition exercise, where we follow the same strategy as for manufacturing and retail. In wholesale and in utilities & transportation, the labor share decline is mild, while the observed increase in sales concentration is also moderate. Consequently, the inferred decline in the price of capital ($d\ln q$) is small, while the inferred increase in competition ($d\ln\lambda$) is weaker than in retail but stronger than in manufacturing. The automation fixed costs (c_a) are small relative to both manufacturing and retail.

TABLE 11: Transitional dynamics and decomposition of the labor share using a non-CES demand system: Wholesale, Utilities & Transportation

		Model		
	Data	Benchmark	Only effects of $d\ln q$	Only effects of $d\ln\lambda$
	(1)	(2)	(3)	(4)
A. Wholesale (1982–2012)				
	I. Parameters and inferred aggregate shocks			
$d\ln q$	.	0.087	0.087	0
$d\ln\lambda$	.	0.240	0	0.240
c_a	.	0.086	0.086	0.086
	II. Targeted moments, 1982–2012			
Δ aggregate labor share	−0.045	−0.045	−0.009	−0.031
Δ log sales concentration	0.202	0.210	0.001	0.207
Δ Relative adoption (P99+ vs. P50-75)	1.71	1.71	2.49	1.88
	III. Other moments, 1982–2012			
Δ median labor share	.	0.023	0.010	0.022
Δ unweighted mean labor share	.	0.016	0.007	0.017
Δ log productivity dispersion	.	0.004	0.000	0.002
	IV. Markups, 1982–2012			
Δ log aggregate markup	.	0.004	0.000	0.004
Within firm change in markups	.	−0.018	−0.007	−0.017
Reallocation	.	0.022	0.007	0.021
B. Utilities & Transportation (1992–2012)				
	I. Parameters and inferred aggregate shocks			
$d\ln q$	.	0.064	0.064	0
$d\ln\lambda$	.	0.150	0	0.150
c_a	.	0.035	0.035	0.035
	II. Targeted moments, 1992–2012			
Δ aggregate labor share	−0.028	−0.028	−0.009	−0.017
Δ log sales concentration	0.108	0.107	0.001	0.105
Δ Relative adoption (P99+ vs. P50-75)	1.71	1.70	2.20	1.94
	III. Other moments, 1992–2012			
Δ median labor share	.	0.011	0.007	0.010
Δ unweighted mean labor share	.	0.007	0.004	0.007
Δ log productivity dispersion	.	0.001	0.000	0.001
	IV. Markups, 1992–2012			
Δ log aggregate markup	.	0.002	0.000	0.002
Within firm change in markups	.	−0.012	−0.007	−0.012
Reallocation	.	0.014	0.007	0.014

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 (BEA-BLS integrated industry-level production account), the relative adoption of automation technologies by firm size (from Acemoglu et al., 2021), and the average log change in the top 4 as well as top 20 firms' sales share (Autor et al., 2020, Table 1), for each sector. 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.

Comparing the various model versions, we find that the declining capital price caused

0.9 pp or 20% of the sectoral labor share decline in wholesale, and 0.9 pp or 32% of the sectoral labor share decline in utilities & transportation. Rising competition accounts for virtually all of the increase in sales concentration, as well as for 3.1 pp or 69% of the sectoral labor share decline in wholesale, and for 1.7 pp or 61% of the sectoral labor share decline in utilities & transportation. The residual labor share decline (around 10% of the overall decline) is due to the interaction 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. In sum, the developments in these sectors are more comparable to retail than to manufacturing.

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 7 provide data on payroll shares by sector for 1947–1987 and 1987–2016 from the BEA industry accounts. We split the data into 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*.

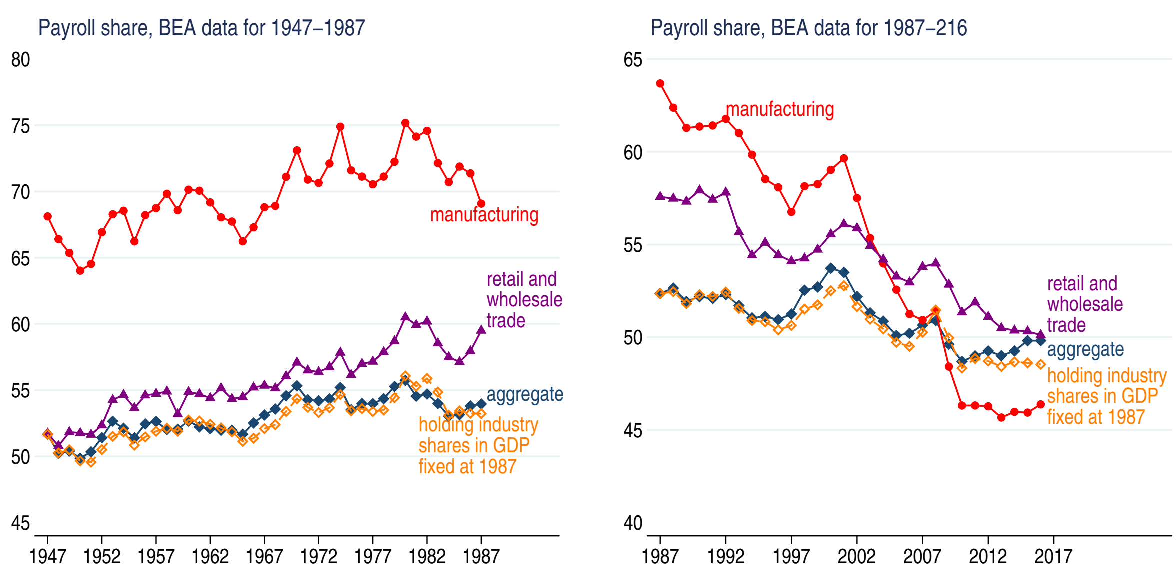

Figure 7: Payroll share in the US for 1947–2016. The figure plots the payroll share of value added, both for some specific sectors and the economy as a whole. Data from the BEA industry accounts. Industry definitions based on SIC in left panel, NAICS in right panel.

As discussed in the main text, Figure 7 shows that payroll shares were constant or increasing up to 1982, and then started a sharp decline both in manufacturing, retail and wholesale. Labor shares (which also include non-wage compensation) are available starting in 1963 from the BEA-BLS integrated industry-level production account. Figure 8 confirms that labor shares, while slightly higher by construction, exhibit the same trend behavior with a flat or slightly increasing trend until 1982 and a subsequent decline. This motivates our focus on the 1982–2012 period and supports our choice of 1982 as the steady state of the model.

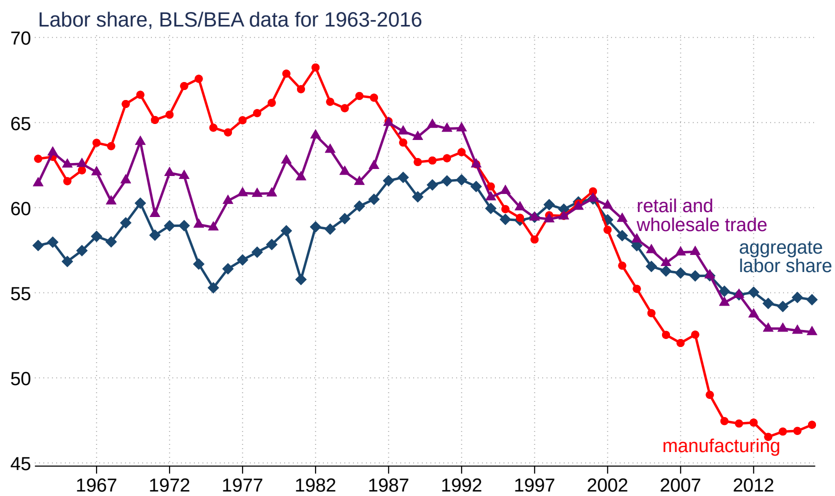

FIGURE 8: LABOR SHARE IN THE US FOR 1963–2016 The figure plots the labor share of value added, both for some specific sectors and the economy as a whole. Data from the BEA-BLS integrated industry-level production account.

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 9 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 for manufacturing and for retail. For simplicity, we focus on the non-CES version of the model calibrated in Section 2.3. As the price of capital declines from 1982 to 2012 and stays constant thereafter, automation cost spending in manufacturing reaches a peak of 1.0% 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 lower, around 0.5% of output in 2005.

The black line provides the behavior of R&D investment 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 45% of manufacturing R&D from 1982 to 2012 correspond to investments in automation fixed costs, and that rising automation cost spending over that time period closely align with the observed 1 percentage point increase in R&D spending. 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.

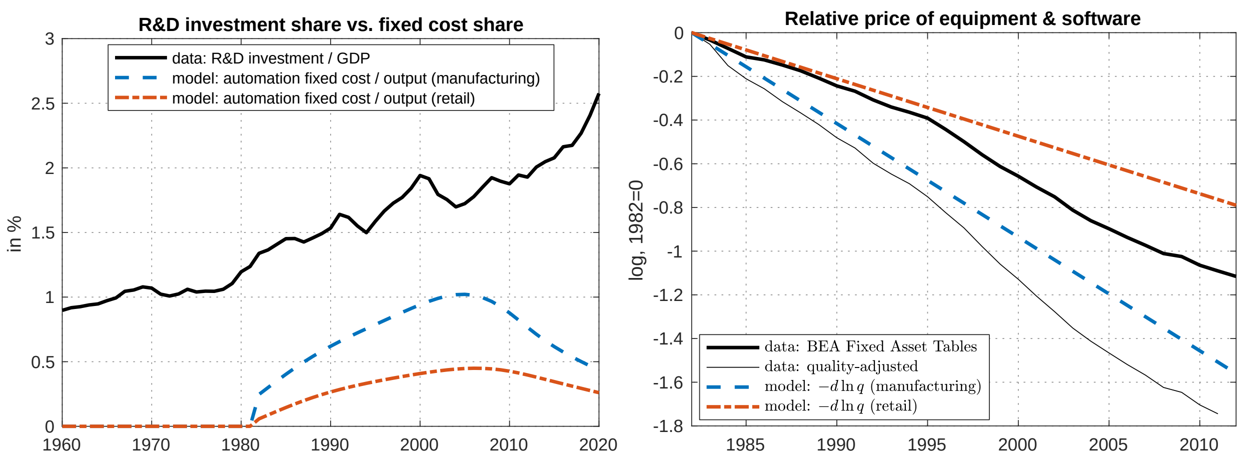

FIGURE 9: AUTOMATION FIXED COSTS AND CAPITAL PRICE DECLINES. The left panel displays the time series of automation fixed cost spending relative to aggregate output for manufacturing and retail using our estimates from Section 2.3. Before 1982, the model is in steady state, and automation spending is zero. The black line corresponds to the ratio of R&D investment spending relative to US GDP. The right panel displays the calibrated decline in capital prices $-d\ln q$ for manufacturing and retail using our estimates from Section 2.3. 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.

Benchmarking the inferred decline in capital prices and rising competition: The right panel of Figure 9 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 equals 156 log points, which is comparable to the observed decline of 174 log points in the data (112 without Gordon's 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 ($d \ln \lambda = 0.30$), followed by wholesale (0.24), utilities/transportation (0.15), and manufacturing (0.05). 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^R_{tf}: we measure revenue using firm sales—*SALES* in Compustat.
- Expenditures in variable inputs v_{tf}: we measure these expenditures using the cost of goods sold— *COGS* in Compustat.
- Stock of capital k_{tf}: we measure capital using the gross value of property, plants, and equipment—*PPEGT* in Compustat.
- Investment rate x_{tf}: we measure the investment rate as the percent change in capital; that is, $\ln x_{tf} = \ln k_{t+1,f} - \ln k_{tf}$
- 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 quin-

tiles, 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 4 to 6.

- Winsorizing: we winsorize the obtained revenue elasticities at zero, and take 5-year moving averages to smooth them. Moreover, following Baqaee and Farhi (2020b), we winsorize our markup estimates at the 5th and 95th percentile of their distribution.

F.2 Estimation approach and details

Given a grouping of firms $c(f)$, we can estimate revenue elasticities following the usual approach from Ackerberg, 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

$$\ln y_{tf}^{R*} = \mathbb{E}[\ln y_{tf}^{R} | \ln x_{tf}, \ln k_{tf}, \ln v_{tf}, t, c(f)] = h(\ln x_{tf}, \ln k_{tf}, \ln v_{tf}; \theta_{tc(f)}^{h}).$$

Here $\theta_{tc(f)}^{h}$ is a parametrization for a flexible function h that might vary over time and between groups of firms. For any pair of revenue elasticities $\varepsilon_{tc(f)}^{Rv}$ and $\varepsilon_{tc(f)}^{Rk}$, one can then compute revenue productivity as

$$z_{tf}^{R} = \ln y_{tf}^{R*} - \varepsilon_{tc(f)}^{Rv} \cdot \ln v_{tf} - \varepsilon_{tc(f)}^{Rk} \cdot \ln k_{tf},$$

estimate the flexible model

$$z_{tf}^{R} = g(z_{t-1,f}^{R}; \theta_{tc(f)}^{g}) + \zeta_{tf},$$

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

$$\mathbb{E}\left[\zeta_{tf} \otimes (\ln k_{tf}, \ln v_{t-1,f})\right] = 0.$$

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

Besides our main estimation approach, we also explored the following variations:

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 10 plots the behavior of the resulting output elasticities over time by firm size quintile. Figure 11 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 revenue productivity directly as

$$z_{tf}^{R} = \ln y_{tf}^{R} - \varepsilon_{tc(f)}^{Rv} \cdot \ln v_{tf} - \varepsilon_{tc(f)}^{Rk} \cdot \ln k_{tf},$$

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

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

$$z_{tf}^{R} = \beta z_{t-1,f}^{R} + \zeta_{tf}.$$

Define $\upsilon_{tf} = z_{tf}^{R} + \epsilon_{tf}$. Because ex-post shocks are i.i.d, we have that υ_{tf} also follows a linear Markov process

$$\upsilon_{tf} = \beta \upsilon_{t-1,f} + \underbrace{\zeta_{tf} + \epsilon_{tf} - \beta \epsilon_{t-1,f}}_{=\iota_{tf}}.$$

Estimation proceeds as follows. First, we can compute υ_{tf} directly as

$$\upsilon_{tf} = \ln y_{tf}^{R} - \varepsilon_{tc(f)}^{Rv} \cdot \ln v_{tf} - \varepsilon_{tc(f)}^{Rk} \cdot \ln k_{tf}.$$

Then we estimate the linear model

$$\upsilon_{tf} = \beta \upsilon_{t-1,f} + \iota_{tf},$$

and base estimation on the moment conditions

$$\mathbb{E}\left[\iota_{tf} \otimes \left(\ln k_{tf}, \ln \upsilon_{t-1,v}\right)\right] = 0.$$

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.

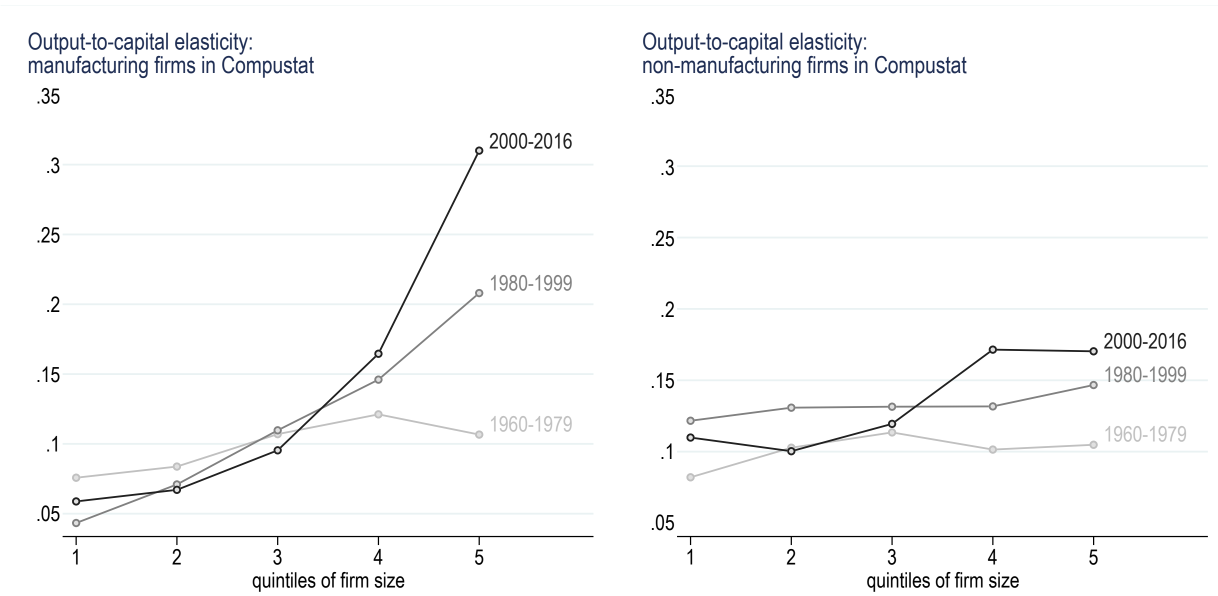

Figure 10: Output-to-capital elasticities for Compustat firms estimated using a cubic parametrization of g and h. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms. Firm-level elasticities are estimated using a cubic parametrization for g and h, as explained in Appendix F. See figure 4 in the main text for our baseline estimates used in the results reported in the paper.

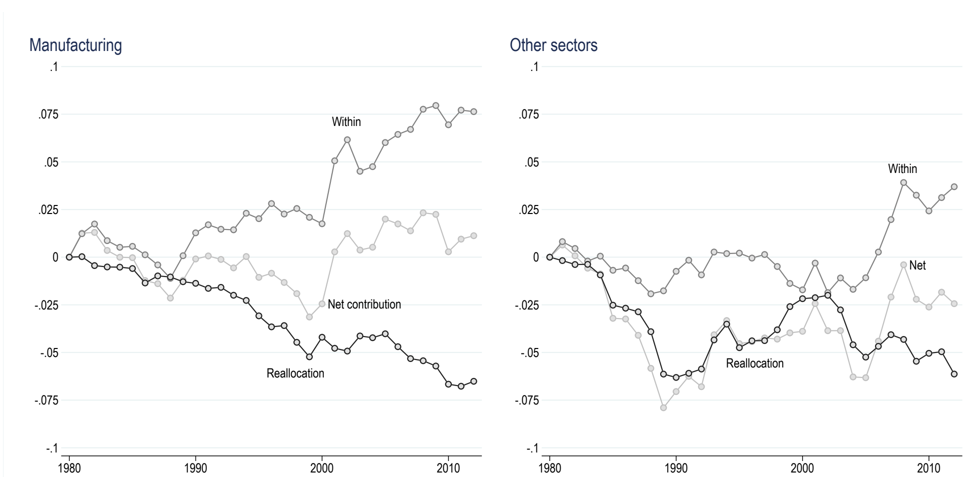

Figure 11: Decomposition of the contribution of within-firm changes in markups and between-firm reallocation to (percent) changes in the labor share. See the main text for details on this decomposition. Firm-level markups are estimated using a cubic parametrization for g and h, as explained in Appendix F. 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 6 in the main text for our baseline estimates reported in the paper.

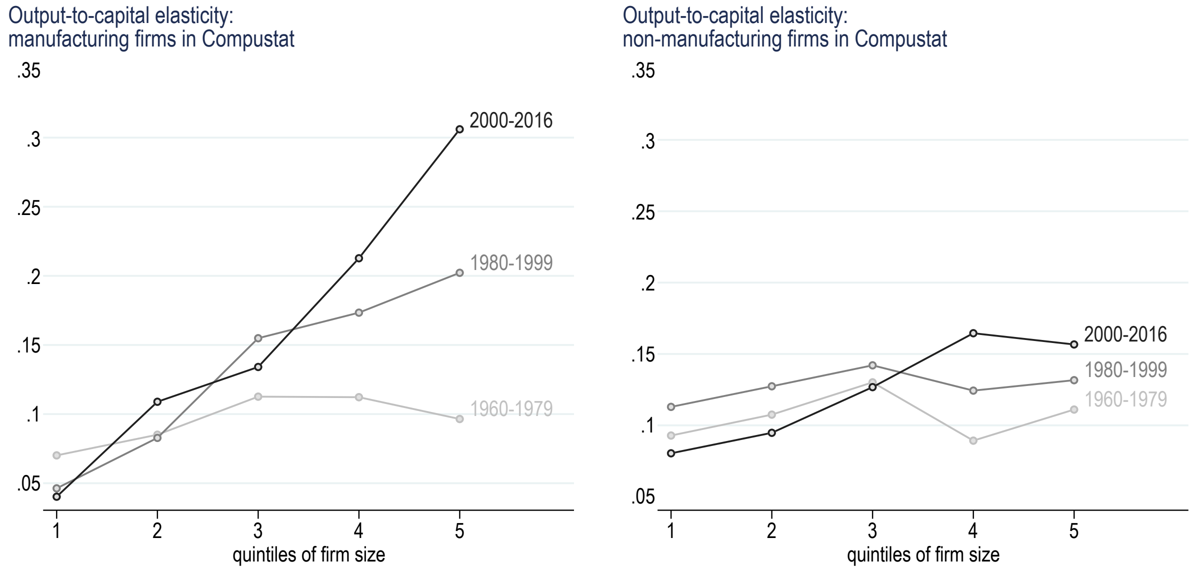

Figure 12: Output-to-capital elasticities for Compustat firms estimated under the assumption that there are no ex-post shocks. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms. Firm-level elasticities are estimated under the assumption of no ex-post shocks, as explained in Appendix F. See figure 4 in the main text for our baseline estimates used in the results reported in the paper.

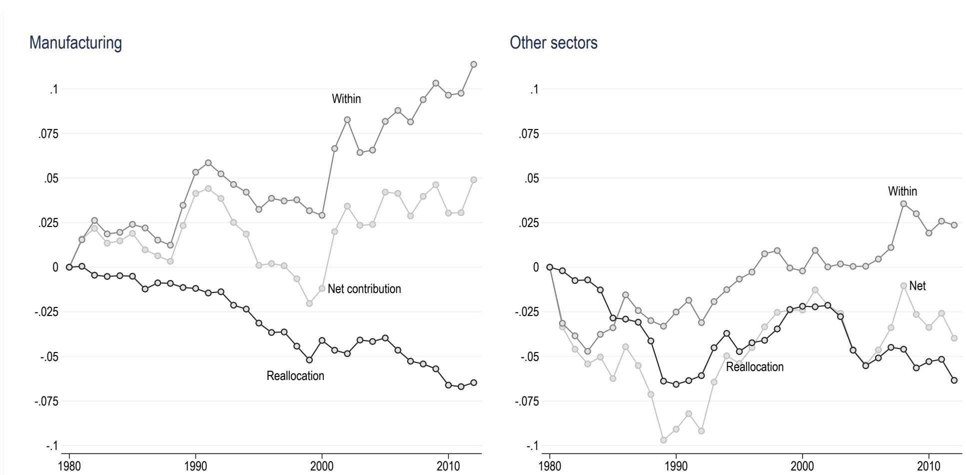

Figure 13: Decomposition of the contribution of within-firm changes in markups and between-firm reallocation to (percent) changes in the labor share. See the main text for details on this decomposition. Firm-level markups are estimated under the assumption of no ex-post shocks, as explained in Appendix F. 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 6 in the main text for our baseline estimates reported in the paper.

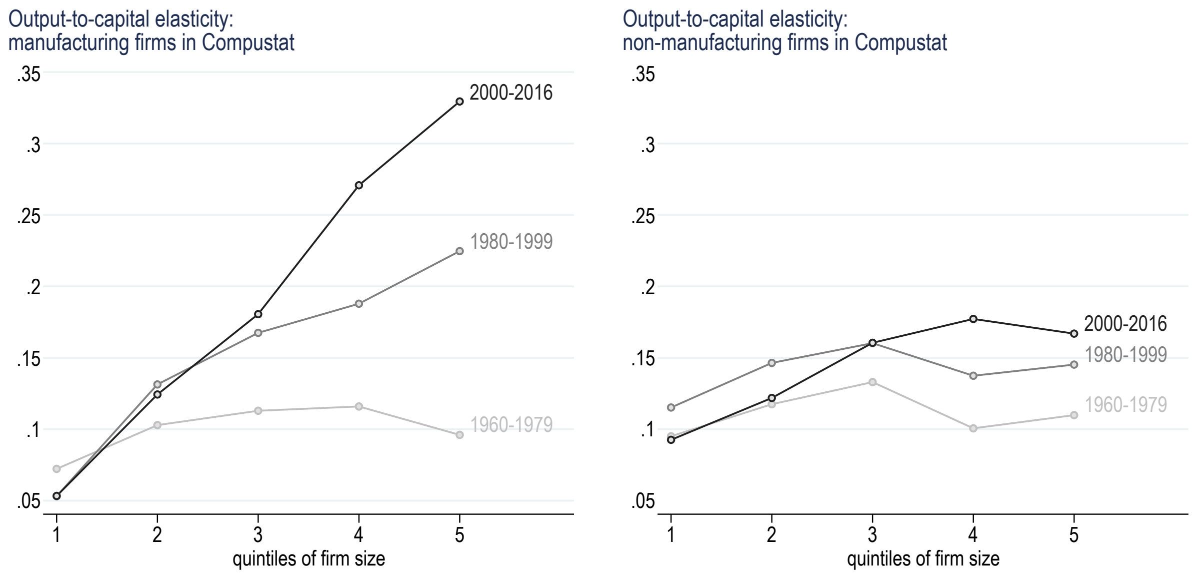

FIGURE 14: OUTPUT-TO-CAPITAL ELASTICITIES FOR COMPUSTAT FIRMS ESTIMATED UNDER THE ASSUMPTION THAT PRODUCTIVITY FOLLOWS A LINEAR MARKOV PROCESS. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms. Firm-level elasticities are estimated under the assumption that productivity follows a linear Markov process, as explained in Appendix F. See figure 4 in the main text for our baseline estimates used in the results reported in the paper.

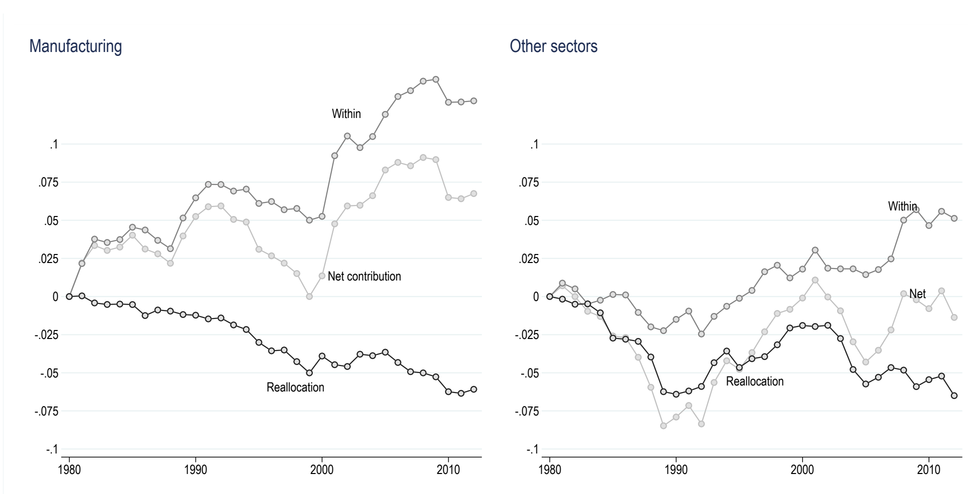

FIGURE 15: DECOMPOSITION OF THE CONTRIBUTION OF WITHIN-FIRM CHANGES IN MARKUPS AND BETWEEN-FIRM REALLOCATION TO (PERCENT) CHANGES IN THE LABOR SHARE. See the main text for details on this decomposition. Firm-level markups are estimated under the assumption that productivity follows a linear Markov process, as explained in Appendix F. 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 6 in the main text for our baseline estimates reported in the paper.