PENN INSTITUTE for ECONOMIC RESEARCH
UNIVERSITY of PENNSYLVANIA

The Ronald O. Perelman Center for Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
23-015

Not a Typical Firm: Capital-Labor Substitution and Firms' Labor Shares

JOACHIM HUBMER
University of Pennsylvania

PASCUAL RESTREPO
Boston University

October 14, 2023

Not a Typical Firm: Capital–Labor Substitution and Firms' Labor Shares*

Joachim Hubmer
jhubmer@sas.upenn.edu
University of Pennsylvania

Pascual Restrepo
pascual@bu.edu
Boston University

October 14, 2023

Abstract

The US labor share has declined, especially in manufacturing and retail. Yet, the labor share of a typical median firm in these sectors has risen. 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 patterns 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 sales concentration in each sector into a part driven by lower capital prices and a part driven by reallocation to higher-markup firms. Lower capital prices played a prominent role in explaining the labor share decline in manufacturing and a smaller role in retail and other sectors from 1982–2012.

*We thank Daron Acemoglu, David Autor, Hugo Hopenhayn, Loukas Karabarbounis, Dirk Krueger, Stephen Terry, and conference and seminar participants at Bank of Mexico, Barcelona GSE Summer Forum, Boston College, Columbia University Junior Conference, Copenhagen, Cornell, Dartmouth, Edinburgh, Essex, Fed Richmond, Glasgow, IIES, IMF, MMF Workshop on Macroeconomic Consequences of Technical Change, NBER Summer Institute, Oslo Macro Conference, Peking University, Philadelphia Fed, SED Minneapolis, Toronto, and Virtual Macro Seminar Series for their comments. Restrepo thanks the National Science Foundation for its support under award No. 2049427.

A striking fact of the recent past is the fall of labor's share of GDP in the United States and other countries.[1] After being stable for much of the last century, the US labor share declined from 62% in the 1980s to 55% in 2012, as seen in Figure 1. The decline is not driven by changes in industry composition: holding industry shares in GDP constant at their 1982 levels leads to a more pronounced labor share decline. As the right panel shows, the fall in the US labor share is driven by a sizable decline dating back to the 1980s of the labor share in retail of 12 percentage points and in manufacturing of 20 percentage points, and a more recent decline since the year 2000 for wholesale, utilities, and transportation.

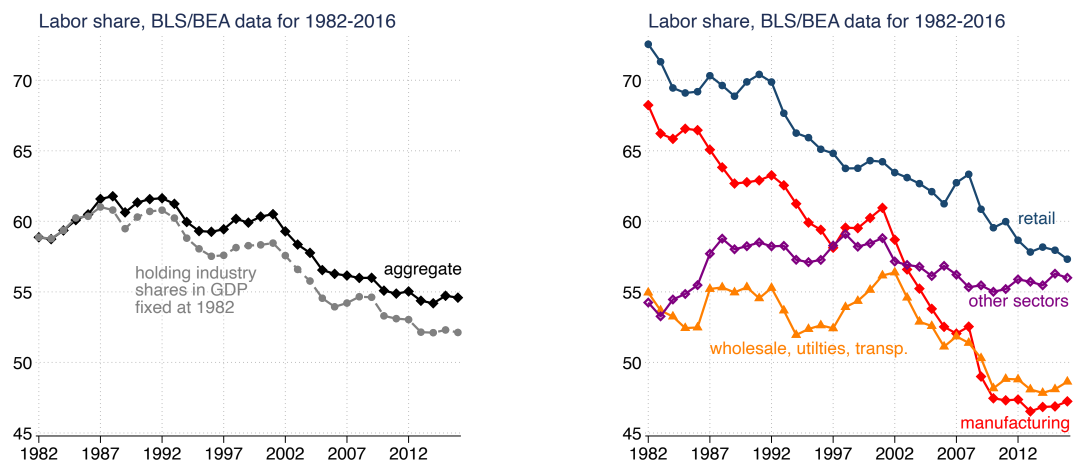

FIGURE 1: US LABOR SHARE. Left panel plots the US labor share for the private nonfarm sector and a counterfactual series holding industry GDP shares constant at 1982 values. Right panel plots labor shares by sector. Data from BEA-BLS integrated industry-level production accounts (Eldridge et al., 2020).

One set of explanations of the labor share decline points to technological advances that favored the substitution of capital for labor. This substitution can take place along an aggregate production function with an elasticity of substitution greater than one (Karabarbounis and Neiman, 2013; Hubmer, 2023) or within tasks as more tasks are automated (Acemoglu and Restrepo, 2018). These explanations receive some support from the fact that the decline in the labor share is more pronounced in manufacturing, and within that sector, in more capital-intensive sub-industries and in firms adopting automation technologies.[2]

This evidence notwithstanding, recent studies show that the decline in labor shares is not

[1] See Elsby, Hobijn and Şahin (2013); Karabarbounis and Neiman (2013); Dao, Das and Koczan (2019).

[2] Acemoglu and Restrepo (2019) document that US industries with greater possibilities for automation saw larger labor share declines. Hubmer (2023) shows that the decline of the labor share is more pronounced in equipment and IT intensive industries. For firms, Acemoglu, Lelarge and Restrepo (2020) and Cheng et al. (2021) document that firms adopting robots in France and China experienced a reduction in their labor shares; Kogan et al. (2021) find that US firms patenting labor-replacing technology see a reduction in their labor shares; and Acemoglu et al. (2022) show that adopters of automation technologies have lower labor shares in all US sectors.

uniform across firms. While the aggregate labor share has declined, the labor share of the *typical* US firm has risen 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, 2021) since the 1980s.

The new firm-level facts cast doubt on explanations of the labor share decline based on capital–labor substitution, since a version of these theories in which firms operate the same technologies implies a uniform decline. The firm-level facts favor a second set of explanations that emphasize the reallocation of sales towards top firms. This reallocation, which could be the result of increased competition or winner-takes-all dynamics, reduces the aggregate labor share because top firms have higher markups and lower labor shares (Autor et al., 2020; Baqaee and Farhi, 2020b; De Loecker, Eeckhout and Unger, 2020).[3]

This paper shows that explanations of the labor share decline based on capital-labor substitution can generate the aggregate and firm-level patterns documented by Autor et al. (2020) and Kehrig and Vincent (2021) if the adoption of capital-intensive technologies to automate additional tasks entails fixed costs. When we account for the observed heterogeneity in technology adoption by calibrating this fixed cost to match available micro evidence, we find that task-based models of capital-labor substitution can account for the aggregate and firm-level facts quantitatively, especially in manufacturing.

As our starting point, Section 1 reviews evidence showing that the adoption of capital-intensive technologies used for automation (i.e., robots, specialized software, and dedicated equipment) concentrates on large firms. Using a new module in the *Annual Business Survey* (ABS) covering all US sectors, Acemoglu et al. (2022) show that the largest firms in an industry are 1.7 times more likely to use automation technologies than the median firm. We complement this evidence with estimates of output elasticities and measures of investment for firms in Compustat. We show that 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.

Motivated by these facts, in Section 2 we build a firm-dynamics model with monopolistic competition and CES demand augmented with costly automation decisions across tasks (as in Zeira, 1998; Acemoglu and Restrepo, 2018). The key innovation is that firms must incur fixed costs to automate tasks, which allows us to match the evidence in Section 1.

We use the model to study the effects of a decline in the *effective price of capital*—or a

[3] A different narrative is that weaker anti-trust eroded competition, raised markups, and led to *bad* concentration (see Philippon, 2019). However, the evidence (including our estimates in Section 3.4) suggests that the markup increase is not due to within-firm changes, but due to reallocation across firms.

q shock—, which captures technological advances increasing the productivity of capital or lowering the cost of producing equipment and software. We show that a q shock can reduce the aggregate labor share at the same time as the median firm labor share rises. The new mechanism driving this result is as follows: in response to a q shock, large and growing firms incur the fixed cost to automate additional tasks, becoming more capital-intensive. In these firms, capital and labor become substitutes, driving the decline of the aggregate labor share. Instead, due to the fixed cost of adoption, the median firm does not automate tasks and continues 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.[4]

Our analytical results show that our model can generate the empirical trends documented in the recent literature. To gauge potential magnitudes, we focus on the manufacturing sector and work with a calibrated version of our model where the fixed cost of automation is calibrated to match the adoption gradient by size in the ABS. We treat the economy in 1982 as being in a steady state—since this period precedes the decline in the manufacturing labor share—and calibrate relevant parameters to match the distribution of firm sales in manufacturing, average markups in the pre-1982 steady state, and other moments. We then use our model to explore the implications of q shocks during 1982–2012.

We consider two q shocks: a *uniform decline* in the price of capital across all tasks and a *skewed decline* that is more pronounced at tasks assigned to labor.[5] The second shock aligns with the definition of automation in Acemoglu and Restrepo (2019) while the first one is reminiscent of the shocks considered in the investment-specific technical change literature (Greenwood, Hercowitz and Krusell, 1997). We show that one can explain the aggregate facts for manufacturing via a combination of q shocks that matches the observed 20 pp decline in the manufacturing labor share during 1982–2012 and is consistent with available estimates of the average decline in capital prices for this period of 108 log points. Our task model generates a decline in the labor share in response to skewed q shocks irrespective of whether the elasticity of substitution between capital and labor is below 1 (as in Oberfield and Raval, 2021) or above 1 (as in Karabarbounis and Neiman, 2013; Hubmer, 2023).

In response to skewed q shocks, the model reproduces the firm labor share moments

[4] A different mechanism introduced by Houthakker (1955) and explored by Jones (2005) emphasizes the possibility that a decline in capital prices reallocates economic activity to more capital-intensive firms, even if firms do not change their factor intensities. Kaymak and Schott (2018) argue that this mechanism generates a third of the manufacturing labor share decline in response to lower corporate taxes. On the other hand, Oberfield and Raval (2021) show that this reallocation has a small effect on the US manufacturing labor share. This form of reallocation operates in our model but is accompanied by an increase in the capital intensity of large firms.

[5] Jones and Liu (2022) and Drozd, Taschereau-Dumouchel and Mendes Tavares (2022) emphasize a similar distinction. Relative to their work, we show that q shocks can generate the labor share dynamics in the data.

observed in the data, even though these are not targeted. The model generates an aggregate labor share decline, while the median and the unweighted mean labor share remain roughly unchanged. As in the Melitz-Polanec decomposition in Autor et al. (2020), the decline in the manufacturing labor share is driven by a more negative covariance between the sales share of firms and their labor share. The more negative covariance reflects the fact that firms automate during periods of expanding sales, which aligns with the dynamics of changes in value added and labor shares for manufacturing firms in Kehrig and Vincent (2021). The uneven use of automation technologies accounts for the increase in TFP dispersion and 75% of the increase in sales concentration in manufacturing from 1982–2012.

Our baseline model in Section 2 abstracts from the role of rising competition in driving the labor share decline. To put both explanations on equal footing, in Section 3 we extend the model to an environment with a non-CES demand system where markups increase with firm size. Following Autor et al. (2020), we model rising competition as resulting from an increase in the mass of customers that a firm can access—a λ shock.[6] This leads to an increase in sales concentration, a reduction in within-firm markups (and an increase in firms' labor shares), and a reallocation towards large firms with high markups (and low labor shares). When the distribution of firm productivity is more log-convex than Pareto, the reallocation effect dominates and the λ shock reduces the labor share.

Both q and λ shocks can generate a pattern of rising concentration, a decline in the aggregate labor share, and a rise in the typical firm labor share. To gauge magnitudes, we use a calibrated version of the non-CES model. For this exercise we also look at retail and other non-manufacturing sectors, where the λ shock might be more prevalent. For each sector, we recover the q and λ shocks from the decline in the sectoral labor share, the average decline in capital prices, and the rise in sales concentration over 1982–2012. This exercise assumes that the observed trends across sectors are driven by a combination of these shocks, recovers the shocks needed to explain the data, and separates their contributions.

One can explain the sectoral trends in labor shares, concentration, and capital prices as resulting from different combinations of q and λ shocks, with these shocks playing different roles across sectors. In manufacturing, skewed q shocks drive most of the decline in the sec-

[6] Previous work has studied other forces driving concentration. Hopenhayn, Neira and Singhania (2022) and Peters and Walsh (2021) study the role of demographics. Aghion et al. (2023); Lashkari, Bauer and Boussard (2019); De Ridder (2020); Hsieh and Rossi-Hansberg (2023) study the role of ICT in allowing more productive firms to expand and create a more contested environment. Our market access shock captures some of these mechanisms in a reduced form way, since a λ shock increases the sales that the most productive firms are able to make but reduces within-firm markups due to tougher competition. In line with Aghion et al. (2023), we find that λ shocks have sizable effects on concentration but mild effects on the aggregate labor share.

toral labor share and explain almost all of the increase in sales concentration, with increased market access playing a small role. Our model with endogenous markups not only matches the aggregate trends but is also consistent with all the new firm-level labor share facts for manufacturing, including the transient behavior of firms' labor shares documented by Kehrig and Vincent (2021) and the labor share decline decompositions in Autor et al. (2020) and Kehrig and Vincent (2021). Instead, in retail, increased market access and the reallocation to high-markup firms accounts for 20% of the decline in the labor share between 1982–2012 and 90% of the increase in sales concentration.

These conclusions align with estimates of firm-level markups from Compustat in Section 3.4. We find that the aggregate manufacturing markup did not increase in this period, but outside of manufacturing, reallocation to high-markup firms reduced the labor share by 6%.

1 The uneven adoption of automation technologies

The most comprehensive evidence on the uneven adoption of automation technologies by large firms comes from the 2019 module of the *Annual Business Survey*, which contains information on the adoption of robotics, dedicated equipment, specialized software, and artificial intelligence for 300,000 firms across all US economic sectors during the period from 2016 to 2018. Using the ABS, Acemoglu et al. (2022) show that, within detailed 6-digit industries, firms in the top percentile of the employment size distribution are 1.71 times more likely to use these technologies than firms between the 50th and 75th percentile. This strong size gradient is visible across sectors and when one conditions on firm age.

Figure 2 reproduces their evidence and plots for each of the technologies in the ABS its adoption rate across firms in different size percentiles (relative to that of firms between the 50th and 75th employment percentiles) within detailed 6-digit industries. Adoption rates at the top are 1.68 (for specialized software) to 3 times higher (for robotics) than in the middle of the firm-size distribution, with an average across technologies of 1.71.

This phenomenon is also visible in other datasets and other countries. Using the *Survey of Manufacturing Technologies* from 1993, Dinlersoz and Wolf (2018) document that the adoption of various automation technologies among US manufacturing firms rises with size. A growing number of papers use data on industrial robot adoption among manufacturing firms in the US, Canada, and several European countries and document that robot use rises sharply with firm size, within narrow industries (see Koch, Manuylov and Smolka, 2021; Humlum, 2019; Bonfiglioli et al., 2020; Acemoglu, Lelarge and Restrepo, 2020). Some of these papers also document that the intensity of use rises with size. For example, using the

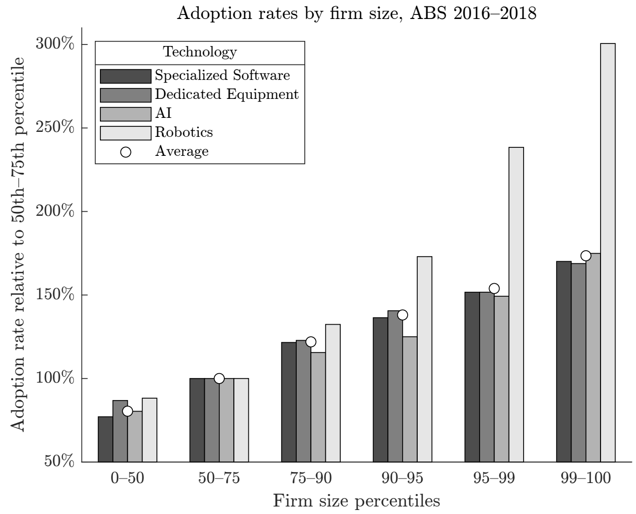

FIGURE 2: ADOPTION RATES BY FIRM SIZE, ABS 2016–2018. The figure reports adoption rates for various capital technologies relative to the adoption of each technology among firms in the 50th–75th percentiles of the employment size distribution. Data from Figure 4 in Acemoglu et al. (2022), reproduced with their permission.

new Annual Survey of Manufactures (ASM), Brynjolfsson et al. (2023) document that the number of robots per worker rises with establishment and firm size.

We complement this evidence with estimates of the output-to-capital elasticity across firms in Compustat. This elasticity—the percent increase in output generated by a percent increase in the quantity of capital—is the key object summarizing the capital intensity of a firm and the extent to which it has automated its production processes. Our production function estimation approach is detailed in Appendix S.1. Relative to the existing literature, we make two modifications. First, we estimate a revenue production function (given the lack of price data in Compustat), but show that one can recover output elasticities by assuming constant returns to scale. Second, we allow output elasticities to vary not only over time and across industries (as in De Loecker, Eeckhout and Unger, 2020), but also across firm-size classes, which accounts for differences in technology across firm sizes.

Figure 3 plots the estimated output-to-capital elasticities by firm size and time period. The left panel shows averaged elasticities for manufacturing industries, while the right panel does so for non-manufacturing. In the 1970s, firms had similar output-to-capital elasticities ranging from 0.08 to 0.12 in both sectors. In the following decades, we estimate a pronounced increase in output-to-capital elasticities among firms in the top quintiles of the size distribution. For the largest manufacturing firms, the elasticity increases by 0.2, going from 0.11 to 0.31. Outside of manufacturing, we estimate a less pronounced increase for large firms of 0.08 (from 0.09 to 0.17). Appendix S.1 confirms that this pattern holds when using descriptive statistics: large firms have higher investment rates and other measures of capital intensity, and these differences have become more pronounced over time.

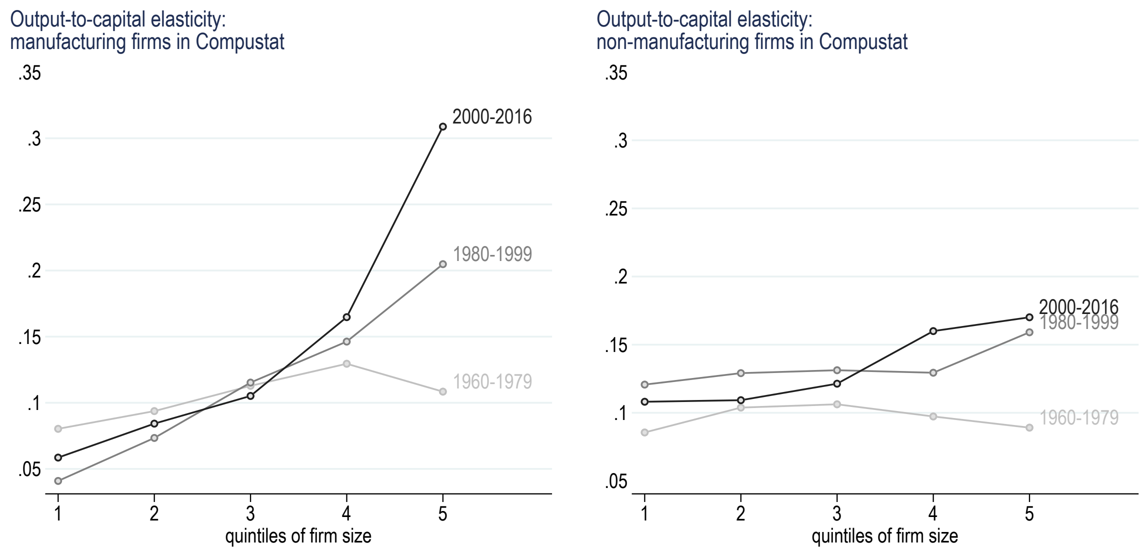

FIGURE 3: OUTPUT-TO-CAPITAL ELASTICITIES FOR COMPUSTAT FIRMS. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms.

We view the evidence from the ABS and Compustat as complementary. The ABS represents the universe of US firms and contains explicit measures of technology adoption, but these measures are only available for 2016–2018 and miss the intensity of use. Compustat contains information on the intensity of capital use and allows us to trace changes in capital intensity over time, but represents a selected sample of firms. While each data source on its own has limitations, together a clear pattern emerges from these sources and the work of Brynjolfsson et al. (2023) and others cited above: large firms differ systematically from small firms in that they are more capital-intensive and have more processes automated.

Our interpretation of these findings is that adopting and integrating automation technologies involves fixed costs.[7] This assumption is plausible. 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 to reorganize its plant, redesign some of their products, and program the robots so that they can be integrated. This is a one-time expenditure that doesn't scale with the quantity of cars produced. The same software routine and plans can be used across factories and once more robots are added to the assembly line. If the firm wants to automate a different process, it must then pay another

[7]There is also direct evidence on these fixed costs. For example, in the case of industrial robots, these upfront integration costs far exceed the cost of the robot system itself (see Acemoglu and Restrepo, 2020).

fixed cost to develop the software and integrate robots to that additional task. Likewise, a firm that wants to deploy a new software to automate its logistics and inventory management must pay a fixed cost for developing the software and rearranging its operations to use it. Here to, the cost of developing the software is independent of firm scale (i.e., the same software system can handle a different number of package deliveries).

In our quantitative exercise, we will use the ABS data in Figure 2 to discipline the size of the fixed costs associated with the adoption and integration of automation technologies. We will treat the estimates from Compustat as *untargeted moments* and explore whether our model can match this rotation in output-to-capital elasticities over time.

2 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 endogenize firms' automation decisions as determined by payment of a fixed cost per task.

2.1 Model and theoretical properties

Environment: 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 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. Non-automated tasks in

$(\alpha_{tf}, 1]$ must be produced by labor. The quantity of task x is then given by

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

where $\ell_{tf}(x)$ and $k_{tf}(x)$ denote labor and capital employed to produce task x with task-specific productivity $\psi(x)$ and $\phi_t(x)$.

Firms compete for a mass ℓ of workers hired 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 transformed into $q_t^I(x)$ units of capital. We treat capital as an intermediate, which is produced instantly and fully depreciates after use. We refer to

$$q_t(x) = q_t^I(x) \cdot \phi_t(x)$$

as the *effective* productivity and $1/q_t(x)$ as the *effective* price of capital and assume $\psi(x)/q_t(x)$ is increasing in x, so labor has comparative advantage at high-index tasks.

Incumbents begin a period with productivity z_{tf} and automation level α_{tf}, make optimal employment and capital utilization decisions, and collect profits π_{tf}. Firms then draw a fixed operating cost $c_o \cdot y_t$, where $c_o \sim G(c_o)$, and decide whether to exit or continue. 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).

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 $c_a \cdot y_t$.[8] 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 draw a productivity signal $z \sim G_e(z)$. After observing z and the realization of the fixed operating cost c_o, entrants decide whether to pay the fixed cost and enter. Firms enter with a common level of automation $\bar{\alpha}_t$, equal to the unweighted average of α_{tf} among incumbents. This is a common specification in models of technology diffusion (see Perla, Tonetti and Waugh, 2021). Diffusion simplifies our analytical results, but is not required for our quantitative results.

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

[8] Fixed costs are scaled by aggregate output to ensure that there is no mechanical relationship between the GDP level in the economy and firms' incentives to produce or automate tasks. The assumption that fixed costs are paid in units of the final good ensures that there is no mechanical relationship between firm size and its labor share.

count the future at a constant interest rate r, which we take as exogenous. 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 capital productivity $q_t(x)$ 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

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

E4. Entry decisions maximize the value function of entrants

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

where z_{tf} denotes an entrant's productivity signal, 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 entry and exit.

Marginal costs and factor shares: The unit cost for firm f at time t is

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

This is the usual CES cost function, with the difference that the *share parameters* $\Gamma^k_t(\alpha_{tf}) = \int_0^{\alpha_{tf}} q_t(x)^{\eta-1} \cdot dx$ and $\Gamma^\ell(\alpha_{tf}) = \int_{\alpha_{tf}}^1 \psi(x)^{\eta-1} \cdot dx$ are endogenous and depend on the degree of automation α_{tf}.

The share of capital in cost ε^k_{tf}—which equals the output-to-capital elasticity—and the share of labor in cost ε^ℓ_{tf}—which equals the output-to-labor elasticity—depend on firm automation decisions, α_{tf}, and are given by

$$\varepsilon^k_{tf} = \frac{\Gamma^k_t(\alpha_{tf})}{\Gamma^k_t(\alpha_{tf}) + \Gamma^\ell(\alpha_{tf}) \cdot w_t^{1-\eta}}, \qquad \varepsilon^\ell_{tf} = \frac{\Gamma^\ell(\alpha_{tf}) \cdot w_t^{1-\eta}}{\Gamma^k_t(\alpha_{tf}) + \Gamma^\ell(\alpha_{tf}) \cdot w_t^{1-\eta}}.$$

Likewise, the labor share *in value added* is $s^\ell_{tf} = \varepsilon^\ell_{tf}/\mu$, which decreases in α_{tf}.

Let α^*_t denote the level of automation that minimizes c_{tf}. This is the point at which the unit cost of producing a task with labor equals that of producing it with capital:

$$\frac{w_t}{\psi(\alpha^*_t)} = \frac{1}{q_t(\alpha^*_t)}. \tag{2}$$

Because of the fixed cost of automation, not all firms automate all tasks up to α^*_t. Proposition 1 shows that, in line with the evidence in Section 1, larger firms automate more tasks. Thus, these firms have higher output-to-capital elasticities ε^k_{tf} and lower labor shares.

Proposition 1 *Assume that for any increasing unbounded function M, $\mathbb{E}[M(z_{t+1,f})|z_{tf}]$ converges to infinity when $z_{tf} \to \infty$ and to $M(0)$ when $z_{tf} \to 0$. Suppose that $\alpha_{tf} < \alpha^*_{t+1}$ for a firm. 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 proposition shows that automation increases with firm size and is episodic. Highly productive (large) firms choose an automation level $\alpha_{t+1,f}$ close to α^*_{t+1}. Moreover, firms go through automation episodes when, following an increase in z, $\tilde{\alpha}_t(\alpha_{tf}, z)$ exceeds α_{tf}. These episodes of automation and declining labor shares coincide with firm expansion.

The effect of a q shock on labor shares and wages: We now explore the effects of q shocks— technological advances that increase the productivity of capital or reduce the cost of capital. We model these advances as a permanent decrease in the effective price of capital $1/q_t(x)$. This can be the result of improvements in the investment technology—an increase in $q^I_t(x)$ reflected in lower capital prices $1/q^I_t(x)$—or advances in the productivity of capital at certain tasks—an increase in $\phi_t(x)$ reflected in a higher quality of capital.

We consider two types of q shocks: a *uniform* decrease in effective capital prices at all tasks, and a *skewed* decrease that is more pronounced for higher-index tasks.

The effects of both shocks can be expressed using two elasticities: the elasticity of substitution between capital and labor holding automation constant, which coincides with the *elasticity of substitution across tasks* η, and the *induced elasticity of substitution*:

$$\eta^*_t = \eta + \frac{\partial \ln[\Gamma^k_t(\alpha)/\Gamma^\ell(\alpha)]}{\partial \ln \alpha} \Bigg/ \frac{\partial \ln[\psi(\alpha)/q_t(\alpha)]}{\partial \ln \alpha}.$$

The induced elasticity accounts for substitution across tasks (given by η) and the automation of additional tasks (given by the change in α_t^*) in response to factor price changes. Because optimal automation decisions are increasing in the wage and decreasing in the effective price of capital (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. This corresponds to the elasticity of substitution that an econometrician exploiting permanent differences in wages across labor markets would estimate. For this reason, we refer to η^* as *the* elasticity of substitution between capital and labor.

The following propositions characterize configurations of η^* and q shocks capable of generating a decline in the aggregate labor share and a simultaneous increase in a typical firm's labor share. We consider an economy that is in a steady state with all firms having a common α^*, and we denote the common cost shares for labor and capital across firms by ε^ℓ and ε^k. We focus on the empirically relevant case with $\eta < 1$, so that tasks are complements.

PROPOSITION 2 (UNIFORM q SHOCK) *Assume tasks are complements ($\eta < 1$). After 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 labor share in cost (or value added) changes by*

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

Along the transition, the labor share of an incumbent firm that automates no additional tasks increases by $d\ln\varepsilon^\ell_{tf} = \varepsilon^k\cdot(1-\eta)\cdot(d\ln w_t + d\ln q) > 0$.

For uniform shocks, the standard argument that the labor share effects depend on the capital–labor elasticity applies. If $\eta^* > 1$, so that capital and labor are substitutes for firms that adjust α optimally, the aggregate labor share declines as capital becomes cheaper.

The second part of the proposition shows that lower capital prices generate dispersed labor share changes along the transition. At one extreme, firms receive positive productivity shocks and reach the scale to justify automating $d\ln\alpha^*$ more tasks. For these firms, capital and labor become substitutes and their labor share falls. At the other extreme, small incumbents keep their α unchanged. The response of these firms to lower capital prices is controlled by the elasticity of substitution across tasks η. For these firms, capital and labor are complements and their labor share rises as capital prices drop and wages increase.

PROPOSITION 3 (SKEWED q SHOCK) *Assume tasks are complements ($\eta < 1$). Following a permanent increase in $q(x)$ of $d\ln q(x) = d\ln q > 0$ for all $x \geq \alpha^*$, 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*

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

which is negative for small $d \ln q$*. Along the transition, the labor share of an incumbent firm that automates no additional tasks increases by* $d \ln \varepsilon^\ell_{tf} = \varepsilon^k \cdot (1 - \eta) \cdot d \ln w_t > 0$*.*

The proposition shows that the effects of (small) skewed shocks on the aggregate labor share are always negative, and this holds irrespective of the elasticity of substitution, η^*.[9] As before, skewed q shocks generate dispersion in labor shares along the transition, with the labor share of some incumbents increasing due to higher wages.

Why do uniform and skewed q shocks differ in their impact? Consider a partial equilibrium example with wages held constant. A uniform q shock involves two effects:

$$d \ln \varepsilon^\ell = \underbrace{\varepsilon^k \cdot (1 - \eta) \cdot d \ln q}_{\text{intensive margin } (+)} + \underbrace{\varepsilon^k \cdot (\eta - \eta^*) \cdot d \ln q}_{\text{extensive margin } (-)}.$$

On the one hand, the reduction in capital prices for tasks above α^* leads to the automation of additional tasks (an increase in α^*). This change at the *extensive margin* 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 change at the intensive margin lowers the share of these capital-intensive tasks in value added and *raises* the labor share (given $\eta < 1$). For a uniform q shock, the extensive margin effect dominates when $\eta^* > 1$. The intensive margin effect dominates when $\eta^* < 1$. Both effects cancel out when $\eta^* = 1$. Instead, a (small) skewed q shock works entirely at the extensive margin and always reduces the labor share.

In general equilibrium, q shocks also increase wages, raising the share in value added of non-automated tasks. In both cases, higher wages increase the labor share of a firm that does not automate additional tasks. This last effect is reminiscent of Baumol's cost disease, as emphasized in Aghion, Jones and Jones (2018), but operating here at the firm level.

Propositions 2 and 3 show that lower capital prices bring a transition period characterized by a declining aggregate labor share, a rising labor share for small incumbents, and an increase in concentration as large and productive firms automate more tasks than small

[9] The expression for $d \ln \varepsilon^\ell$ shows that skewed q shocks also affect the labor share via wages in general equilibrium. By definition, the effect of wages on the aggregate labor share is mediated by the elasticity of substitution η^* and could be positive or negative. The proof of the proposition shows that this effect via wages is second order.

incumbents. This outcome can be explained in two ways: (i) by having a uniform q shock and $\eta^* > 1$, as emphasized in Karabarbounis and Neiman (2013), Hubmer (2023) and in Proposition 2; or (ii) by having a skewed q shock, as emphasized in Proposition 3. 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 (2021).

In our quantitative exercise we focus on (ii) and allow for both skewed and uniform q shocks. This choice allows us to explore scenarios with values of the elasticity of substitution below or above 1, and emphasizes the novel fact that our explanation for the labor share decline does not require an elasticity of substitution above 1.[10]

Besides matching the data for a wide range of parameters, skewed and uniform q shocks capture distinct forms of technological progress that are relevant in practice.[11] Skewed q shocks align with the definition of automation in Acemoglu and Restrepo (2019). These correspond to advances in the creation of new types of equipment and software capable of producing more complex tasks that were previously the domain of human labor. For example, we think of many recent advances in robotics, computer-assisted manufacturing, and artificial intelligence as skewed q shocks that work primarily at the extensive margin. Industrial robots are capable of performing tasks such as welding and assembly that previous machines could not perform. Likewise, large-language models (LLMs) and related AI models work by expanding the range of tasks that we can automate via software systems.

In contrast, uniform q shocks correspond to the traditional definition of investment-specific technical change (Greenwood, Hercowitz and Krusell, 1997). They capture advances in capital that occur mostly at the intensive margin (i.e., at tasks below α^*). For example, improvements in farm equipment and tractors take place at the intensive margin since these tasks have long been mechanized. Other examples are more efficient industrial furnaces or boilers, improvements in electric generators and transmission equipment, and cheaper vintages of an existing machine.

[10] As explained in footnote 2, recent capital advances have been shown to be associated with declining labor shares in multiple contexts (e.g., Karabarbounis and Neiman (2013) and Hubmer (2023)). Our theory shows that this is possible even if the capital–labor elasticity is < 1. In this case, these studies can be reinterpreted as providing evidence of recent advances in capital being skewed towards more complex tasks. This interpretation reconciles these findings with the work of Oberfield and Raval (2021) and others, who find limited substitution between capital and labor (i.e., $\eta^* < 1$) in response to wage differences across regions (akin to a uniform q shock).

[11] A growing literature recognizes the distinction between uniform and skewed q shocks and the extensive and intensive margin of advances in capital. Acemoglu and Restrepo (2019) distinguished automation—advances in capital at the extensive margin (as in our skewed q shock), from a deepening of capital—advances at the intensive margin. Jones and Liu (2022) also emphasize this distinction and show that one can generate balanced growth from a combination of extensive margin advances and intensive margin improvements in capital.

2.2 Calibration and quantitative results

This subsection shows that a calibrated version of our model reproduces the firm-level patterns in US manufacturing in response to q shocks of a magnitude that aligns with available data on capital prices.

We take the economy in 1982 to be in steady state with all firms having the same level of automation α_0. We calibrate the parameters of our model to match key moments of the US manufacturing sector in the 1982 steady state and study a reduction in effective capital prices from 1982 to 2012. The timing is motivated by the fact that the decline of the manufacturing labor share starts in 1982, after being constant in the decades preceding it.[12]

Calibration of initial steady state: A combination of a uniform and a skewed q shock can generate a decline in the manufacturing labor share even for an elasticity of substitution η^* less than or equal to one. To demonstrate this novel aspect of our model, we consider a parametrization with $\eta^* = 1$ as our baseline and report results for a calibration with $\eta^* > 1$ at the end of this section. We normalize initial capital prices by task $q_t^I(x)$ to 1 and parametrize capital and labor productivity across tasks as:

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

where $\gamma_k, \gamma_\ell > 0$ control the strength of the comparative advantage of labor at higher indexed tasks. 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

$$y_f = z_f \cdot \left(\gamma_k^{-\frac{1}{\eta}} \cdot \left(\frac{\alpha_f}{1-\alpha_f} \right)^{\frac{\gamma_k}{\eta}} \cdot k_f^{\frac{\eta-1}{\eta}} + \gamma_\ell^{-\frac{1}{\eta}} \cdot \left(\frac{\alpha_f}{1-\alpha_f} \right)^{-\frac{\gamma_\ell}{\eta}} \cdot (A \cdot \ell_f)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} . \tag{3}$$

Maximizing (3) with respect to α yields the induced production function

$$y_f^* = z_f \cdot \mathcal{A} \cdot k_f^{\frac{\gamma_\ell}{\gamma_k+\gamma_\ell}} \cdot \ell_f^{\frac{\gamma_k}{\gamma_k+\gamma_\ell}},$$

where $\mathcal{A}$ is a constant. This shows that, in the initial steady, the induced elasticity of substitution equals 1 and all firms operate a Cobb-Douglas production function with share parameters $\gamma_\ell/(\gamma_k+\gamma_\ell)$ and $\gamma_k/(\gamma_k+\gamma_\ell)$ for capital and labor, respectively.

[12] Appendix S.2 shows that this also holds for other sectors. After having stable labor shares during 1947–1982, some sectors experienced a decline in their labor shares starting in the mid 1980s. Available measures of capital prices also point to a more rapid decline starting in the 1980s (see Hubmer, 2023). Moreover, as shown in Section 1, firms of different sizes in had similar capital intensities before 1980, which points to comparable α_0's.

We pick $\mathcal{A}$ to normalize wages in the 1982 steady state to 1. Initial factor shares depend only on γ_ℓ/γ_k; we normalize $\gamma_k = 1$ and set $\gamma_\ell = 0.3$ to match an initial manufacturing labor share of 67%. We also set $\eta = 0.5$ to match the task-level elasticity of substitution estimate in Humlum (2019). This parametrization is consistent with papers that estimate elasticities of substitution between capital and labor below 1. For example, exploiting differences in wages across regions, Oberfield and Raval (2021) estimate elasticities of substitution in the 0.5–1 range. In our model, these estimates should lie between η—for firms that have not adjusted their α_f's—and 1—for firms that set α_f optimally in response to wage changes.

We calibrate the demand system, productivity process, and operating costs to match moments from the US manufacturing sector in 1982. 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 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 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 (based on the number of firms per industry reported by Autor et al.). We calibrate the fixed cost of operation and entry (and their dispersion) to match entry rates, exit rates, and the relative size of exiters and entrants reported in Lee and Mukoyama (2015).[13] 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 from the production function estimates in Section 1. This aligns with estimates using US Census data in Foster, Haltiwanger and Syverson (2008).

Decline in the price of capital and fixed cost of automation: We study the effects of a decline in capital prices given by a combination of uniform and skewed q shocks. We model these as advances in the productivity of the investment sector, $q_{I,t}(x)$, varying across tasks:

$$q_{I,t}(x) = \begin{cases} q_{u,t} & \text{if } x \leq \alpha_0^* \\ q_{u,t} \cdot \min\{\psi(x)/\phi(x), q_{s,t}\} & \text{if } x > \alpha_0^* \end{cases}$$

[13] 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 with a lower mean $\mu_e < 1$.

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_ℓ	Comparative advantage	0.30	Manufacturing labor share (BLS/BEA)	0.67	0.67
	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 1)		

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 term $q_{u,t}$ captures the uniform component of the q shock and $q_{s,t}$ captures the skewed component. The formulation of the skewed component mimics a sequence of small skewed shocks as the one considered in Proposition 3. $q_{u,t}$ and $q_{s,t}$ increase log-linearly from a level of 1 in 1982 to new steady state levels $\bar{q}_u$ and $\bar{q}_s$ in 2012, with $d\ln q_u = \ln \bar{q}_u > 0$ and $d\ln q_s = \ln \bar{q}_s > 0$ summarizing these improvements in investment technologies.

To discipline the extent of technological progress in the production of capital goods, captured by $d\ln q_u$ and $d\ln q_s$, we use data on investment-good prices from the BEA and adjust it for quality improvements using the methodology of Cummins and Violante (2002), extended to recent years by DiCecio (2009).[14] This approach yields an average price decline for capital goods in manufacturing of 108 log points.

[14] Cummins and Violante (2002) estimate a statistical model explaining Gordon (1990)'s price indices as a function of those by the BEA/NIPA, their lags, and time trends. They extrapolate this model to produce quality-adjusted price indices for 1947–2000. We use the updated estimates from DiCecio (2009) to compute a quality adjustment for equipment and software of 2.6% per year for 1980–2011. We then compute the decline in investment-good prices for manufacturing using a user-cost weighted Törnqvist index of the prices for equipment and software (quality adjusted) and structures from the BEA Fixed Asset Tables. We account for changes in effective capital taxes from Acemoglu, Manera and Restrepo (2020). Appendix S.2 provides details.

We recover $d\ln q_u$ and $d\ln q_s$ from: (i) the *average* decline in quality-adjusted capital prices in manufacturing of 108 log points; and (ii) the 20 pp decline in the manufacturing labor share observed from 1982 to 2012. Intuitively, a larger labor share decline requires a q shock that loads more on the skewed component $d\ln q_s$. On the other hand, a large drop in capital prices that is not associated with a large labor share decline signals a uniform q shock.[15] Our procedure ensures that the total amount of technological improvements in capital are consistent with the observed decline in capital prices, and then shows that one can match the labor share decline via a combination of skewed and uniform q shocks.

In addition, we calibrate the fixed cost of automating tasks required to match the uneven adoption of automation technologies in the ABS. As documented by Acemoglu et al. (2022), firms in the top percentile of the employment distribution are 1.71 times more likely to have adopted new automation technologies than firms between the 50th and 75th percentile. We calibrate c_a so that by 2016–2018

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

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

Panel I in Table 2 reports the resulting shocks and fixed automation cost needed to match the manufacturing data. A uniform decrease in effective capital prices of $d\ln q_u = 0.67$ and a greater decrease of $d\ln q_s = 4.81$ skewed towards complex tasks match (i) and (ii). In our calibration, the decline in capital prices occurs mostly at the extensive margin, explaining the large drop in the manufacturing labor share. This is in line with evidence on the widespread use of capital-intensive technologies for automation in manufacturing.[16]

The fixed cost of automation of 0.19 required to match the adoption rates in the ABS data is of reasonable magnitude. This cost can be thought of as an investment in R&D required to design and integrate automation equipment or software. In response to lower capital prices, aggregate spending on automation fixed costs as a share of manufacturing output reaches a peak of 1.9% by 2005. This is a fifth of total R&D expenditure in manufacturing during this period, which is reassuring since not all R&D expenditure is linked to automation.

Lower capital prices and concentration: Although not targeted, our model generates most of the observed increase in sales concentration and productivity dispersion in manufacturing

[15] Observed price series for capital are informative of *average* advancements in capital but not of the skewness of the q shock, precluding direct measurement of $d\ln q_u$ and $d\ln q_s$.

[16] See, for example, Acemoglu and Restrepo (2020) and the direct survey evidence in Acemoglu et al. (2022).

TABLE 2: Labor share dynamics in manufacturing using the CES demand system (1982–2012)

	Data	Model: Baseline CES model	Model: No automation fixed cost	Model: No technology diffusion	Model: Induced elasticity $\eta^* = 1.45$
	(1)	(2)	(3)	(4)	(5)
I. Parameters and inferred aggregate shocks					
$d\ln q_u$		0.67	0.67	0.67	0.86
$d\ln q_s$		4.81	4.81	4.75	0.77
c_a		0.19	0	0.09	0.32
II. Targeted moments, 1982–2012					
Δ aggregate labor share	-0.20	-0.20	-0.20	-0.20	-0.20
Δ log average capital price	-1.08	-1.08	-1.10	-1.08	-1.09
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1.00	1.69	1.70
III. Concentration 1982–2012 (from Autor et al., 2020; Decker et al., 2020)					
Δ log 4 firms' sales share	0.140	0.105	-0.022	0.059	0.055
Δ log 20 firms' sales share	0.072	0.104	-0.017	0.062	0.054
Δ log productivity dispersion (from 1980–2000)	0.050	0.061	-0.001	0.085	0.041
IV. Typical firm labor share from Kehrig and Vincent (2021), 1982–2012					
Δ median labor share	0.030	-0.005	-0.202	0.085	-0.042
Δ unweighted mean	-0.017	-0.039	-0.203	-0.026	-0.055
V. Melitz–Polanec decomposition from Autor et al. (2020)					
Δ aggregate labor share	-0.185	-0.198	-0.202	-0.198	-0.202
Δ unweighted incumbent mean	-0.002	-0.015	-0.202	-0.062	-0.029
Exit	-0.055	-0.006	0.000	-0.003	-0.004
Entry	0.059	0.006	0.000	0.004	0.006
Covariance term	-0.187	-0.183	0.000	-0.138	-0.174
VI. Covariance decomposition from Kehrig and Vincent (2021)					
Market share dynamics	0.047	0	0	0	0
Labor share by size dynamics	-0.043	-0.134	-0.202	-0.155	-0.123
Cross-cross dynamics	-0.232	-0.066	0	-0.044	-0.083

Notes: Column (2) reports the findings from our benchmark model. Column (3) displays a counterfactual economy with no fixed cost of automation. Column (4) displays a re-calibrated economy with no diffusion through entry. Column (5) features a calibration with an induced elasticity of substitution above one ($\eta^* = 1.45$).

since 1982 in response to the calibrated q shocks. As shown in Panel III of Table 2, by 2012, the uneven adoption of automation technologies by the most productive firms leads to a 10.5 log points (4.4 pp) increase in the share of sales among the top 1.1% firms, and a 10.4 log points (7.6 pp) increase in the share of sales among the top 5.5% firms in manufacturing. 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).[17] The model also generates an increase in productivity (TFP) dispersion across firms of 6.1 log points over 1980–2000, which matches the 5 log points increase estimated by Decker et al.

[17] As in the data, the model generates a larger increase in sales than in employment concentration.

(2020) using US manufacturing data for this period.

Lower capital prices and firms' labor shares: Although not targeted, our model matches several firm-level labor share facts. Column 2 of Panel IV show that our model generates the 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 3.9 pp—comparable to the data, where it decreased by 1.7 pp.

To illustrate this point, Figure 4 depicts the evolution of the labor share across the distribution of firm productivity z_{tf}. The lines trace the average 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 3

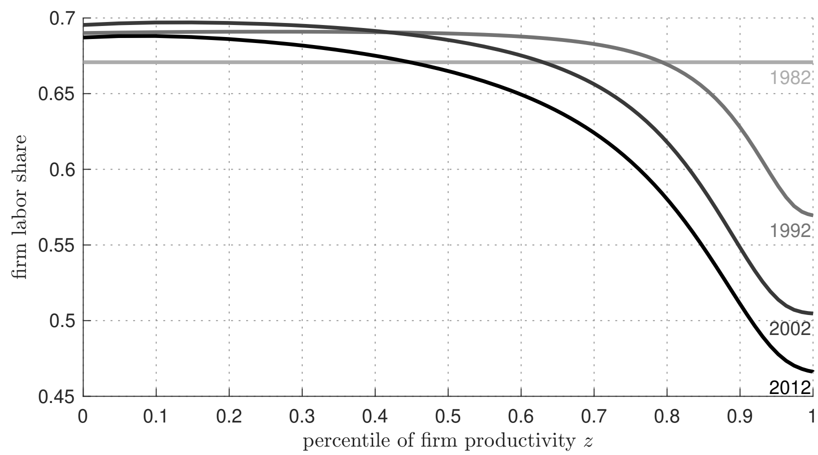

FIGURE 4: CROSS-SECTIONAL LABOR SHARES OVER THE TRANSITION. Firm labor shares in value added by firm productivity z_{tf} from the benchmark model with CES demand.

This clockwise rotation in labor shares at the top coincides with the counter-clockwise increase in output-to-capital elasticities documented for Compustat firms in Section 1. In fact, our model generates an increase in the output-to-capital elasticity for large firms in manufacturing of 20 pp, in line with the 20 pp increase estimated in Compustat.

Our model also generates the firm-level labor share patterns documented by Autor et al. (2020) and Kehrig and Vincent (2021) as features of the transition of an economy in response to q shocks. Autor et al. decompose the decline in the manufacturing labor share using a

Melitz–Polanec decomposition:

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

Here, Δs_t^ℓ denotes the change in the manufacturing labor share between two periods, t and t', corresponding to five-year periods. This can be decomposed into the change in the unweighted mean of labor shares among continuing firms, $\Delta \bar{s}_t^\ell$; two terms accounting for the contributions of exit and entry; and the change in the covariance among continuing firms between their share of value added, ω_{tf}, and their labor share, s_{tf}^ℓ.[18]

We conduct the same decomposition in our model for 5-year differences spanning the 1982–2012 period. Panel V in Table 2 reproduces Autor et al.'s manufacturing data and reports the decomposition from our model. In 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.[19]

Kehrig and Vincent show that the covariance term can be further decomposed into three terms capturing the joint dynamics of firms' labor shares and sales:

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

A decrease in the covariance can be driven by a reallocation of value added towards firms with lower labor shares at baseline (the "market share dynamics" term); a more pronounced reduction in the labor share of large firms (the "labor share by size dynamics" term); or the possibility that firms that reduce their labor shares expand at the same time (the "cross-cross dynamics" term). Using a balanced sample of firms for 1982–2012, Kehrig and Vincent

[18] The contribution of exit equals the value added share of exiting firms, ω_{tX}, times the difference in the mean labor share of continuing firms, s_{tS}^ℓ, and firms that exit, s_{tX}^ℓ. The contribution of entrants equals their value added share, $\omega_{t'E}$, times the difference in the mean labor share of firms that enter, $s_{t'E}^\ell$, and continuing firms, $s_{t'S}^\ell$.

[19] Exit and entry exhibit the same qualitative patterns as in the data, with both entering and exiting firms having higher labor shares than 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 elements of firms' life cycle not in our model.

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 increased the manufacturing labor share by 4.7 pp.

Panel VI in Table 2 compares the contributions of these components in data and model.[20] In our model, the labor share dynamics by size contributed a 13.4 pp decline to the labor share, and the cross-cross dynamics contributed a 6.6 pp decline to the labor share (by construction, the market share dynamics is zero since all firms had the same labor shares in the 1982 steady state). The reason why the cross-cross dynamics contribute to the decline of the labor share in our model is that firms automate tasks following high productivity draws. Because these firms simultaneously gain market share and reduce their labor shares, this shows up as part of the covariance term in the Melitz–Polanec decomposition in Autor et al. and as part of the cross–cross term in the decomposition in Kehrig and Vincent. Although our model with CES demand and constant markups matches the covariance decomposition qualitatively, it does not generate the large cross-cross term seen in the data. Our model with endogenous markups in the next section solves this problem and generates cross-cross dynamics comparable to the data.

Additional specifications: Columns 3–5 in Table 2 report variants of our calibration. Column 3 summarizes the transitional dynamics of our model in response to the calibrated q shocks but now assuming that firms faced no fixed costs of automating tasks. The aggregate labor share evolves similarly in this economy, but the firm-level behavior of labor shares are at odds with the data. The comparison between Columns 2 and 3 shows that a small fixed cost of automating tasks is needed to make sense of the manufacturing firm-level data.

Column 4 shuts down the diffusion of automation through entry (so that entrants start with the same level of automation as in the initial steady state) and re-calibrates the q shocks and the fixed cost of automation required to match the manufacturing data. We find similar results to Column 2. This shows that diffusion is not crucial for our findings for 1982–2012. Its main role is to facilitate characterizing the steady state of the model.

Column 5 explores the implications of using a calibration with an elasticity of substitution above 1. Our model can also match the facts if one adopts a parametrization with an induced elasticity of substitution of $\eta^* = 1.45$, as estimated by (see Karabarbounis and Neiman, 2013; Hubmer, 2023). Appendix A.3 provides the details of this parametrization. As expected, the q shock required to explain the data is more uniform and less skewed, since the uniform

[20] We compute the model moments as in Kehrig and Vincent (2021) as the 30 years change in a balanced sample. Thus, in model and data, the terms in Panel V do not exactly add up to the covariance term in Panel IV.

shock now generates a large reduction in the labor share. This calibration also matches the aggregate and firm-level labor share patterns in manufacturing, though it generates a larger decline in the median and unweighted mean of firms' labor shares.

Taking stock: Our approach assumes that the decline in the manufacturing labor share is driven by q shocks, and infers their extent to match the manufacturing labor share behavior. Because of this, the results in this section are *possibility results*. They show that it is possible to have a coherent description of the manufacturing labor share decline driven by lower capital prices that: (i) fits both the macro and micro patterns of the labor share decline; (ii) is consistent with the available data on capital prices; and (iii) does not require an elasticity of substitution between labor and capital greater than one. Our exercise also shows that a small fixed cost of automating tasks is necessary and sufficient to make sense of the manufacturing firm data and generate most of the increase in sales concentration in the sector as a response to lower capital prices.

3 Capital–labor substitution vs. rising competition

This section extends our model to allow for endogenous firms' markups. This allows us to match the decline in the labor share and the rise in sales concentration and decompose them into a part driven by rising competition and another driven by advances in capital.

3.1 Model extension

Environment: We retain the production side of the model in section 2 and modify the demand system to allow for endogenous markups. We work with a model of monopolistic competition in which firms produce differentiated varieties and face a demand derived from a Kimball aggregator (Kimball, 1995). There is a mass 1 of customers indexed by j whose flow utility u_{tj} from consuming a set of varieties $\mathcal{F}_{tj}$ is defined implicitly by

$$\int_{f\in\mathcal{F}_{tj}} H\left(\frac{y_{tfj}}{u_{tj}}\right)\cdot df = 1,$$

where H is an increasing and concave function.

Firms and customers are randomly matched to each other, with each customer matched to a mass λ_t of firms. This implies that total demand for firm f is

$$y_{tf} = y_t \cdot \lambda_t \cdot D\left(\frac{p_{tf}}{\rho_t}\right), \tag{4}$$

where y_t denotes total expenditure in the economy. Here, D is decreasing and given by the inverse function of $H'(x)$, and ρ_t is an endogenous summary measure of competitors' prices, which is common to all customers and is defined implicitly by

$$\lambda_t \cdot \int_f H\left(D\left(\frac{p_{tf}}{\rho_t}\right)\right) \cdot df = 1. \tag{5}$$

Adding (4) up across firms we get $y_t = \int_f y_{tf} \cdot p_{tf} \cdot df$, which implies that firm prices satisfy the ideal-price index condition

$$1 = \lambda_t \cdot \int_f p_{tf} \cdot D\left(\frac{p_{tf}}{\rho_t}\right) \cdot df. \tag{6}$$

This section considers the effects of q shocks and an exogenous increases in λ_t—or λ shocks. This captures an increase in *market access*: each firm now competes against more firms for each customer. For example, one might think of advances in online marketing and platforms as allowing each firm to reach a larger set of customers (see, for example, Akerman, Leuven and Mogstad, 2022). Or advances in IT allowing firms to enter into additional markets at a lower overhead cost, raising competition for customers in those markets (see, for example, Aghion et al., 2023; Hsieh and Rossi-Hansberg, 2023).

The effects of a λ shock depend on the curvature of demand. We assume that demand satisfies Marshall's weak (demand elasticity $-x \cdot D'(x)/D(x) > 1$ and increasing in x) and strong (marginal revenue $x + D(x)/D'(x) > 0$ and log-concave) second laws.[21] Marshall's weak second law requires that, as firms lower their prices, their demand becomes more inelastic. This implies that more productive firms charge lower prices *and* higher markups. The strong second law implies lower passthroughs for more productive firms. These assumptions ensure that more productive firms charge higher markups, and that an increase in competition will reallocate economic activity towards these high-markup firms.[22]

Equilibrium: Given a path for effective capital prices $1/q_t(x)$ and market access λ_t, and an initial distribution of firms, an equilibrium is defined as before. The difference is that now sales and profits depend on the demand curve D in (4) and we also have to solve for the competitors' price index ρ_t using (5) and for wages using the price condition in (6).

Effects of rising competition: We now derive the steady-state effects of an increase in

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

[22] These assumptions receive empirical support (see Baqaee and Farhi, 2020a) and capture the pricing dynamics of oligopolistic competition models, where large firms face more inelastic demand. Importantly, these assumptions capture the key mechanism in theories that see rising competition as the driving force behind the labor share decline. For example, Autor et al. (2020) use this model to illustrate how competition can reduce the labor share.

market access. As in our baseline model, the equilibrium of the economy converges to a steady state where $\alpha_{tf} = \alpha^*$. Because of this, 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 4 *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;*
- *for $z > z'$, $\mu_z/\mu_{z'}$ decreases;*
- *for $z > z'$, $\omega_z/\omega_{z'}$ increases.*

Market access strengthens firms competition for workers, which leads to higher real wages. This then pushes firms towards the more elastic segments of their demand curves and firms respond by reducing their markups (the weak second law). The reduction in markups is not uniform. Because large firms have smaller passthroughs (the strong second law), they respond via a modest increase in their prices (and a large reduction in their markups). Small firms respond via a more sizable increase in their prices (and a smaller reduction in their markups). As a result, the λ shock reallocates economic activity and labor towards the largest and most productive firms in the industry.

These responses generate an ambiguous contribution of markups to the aggregate labor share. Firms reduce their markups, which contributes to an increase in the aggregate labor share. But the reallocation of economic activity from small low-markup firms to large high-markup firms contributes to a decline in the aggregate labor share.

Proposition 5 *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).[23]

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

The proposition shows that the effect of market access on the labor share via markups depend on the distribution of firm productivity. This insight is well known and is relevant for understanding the effects of a market access shock on the labor share.

3.2 Calibration for manufacturing and retail

We now explore the adjustment of the economy in response to rising market access and lower capital prices in a calibrated version of our model.

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

Relative to the previous section, we make two modifications. First, following Edmond, Midrigan and Xu (2022), we parametrize H using the specification from Klenow and Willis (2016), which implies that the demand elasticity faced by a firm with price p_{tf} is

$$\text{demand elasticity}(p_{tf}) = \sigma \cdot D\left(\frac{p_{tf}}{\rho_t}\right)^{-\frac{\nu}{\sigma}}. \tag{7}$$

This demand elasticity decreases as p_{tf} falls, so that more productive firms face a more inelastic demand.[24] 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. (If $\nu = 0$, the demand system simplifies to the standard CES aggregator.)

Second, we adopt a new process for firm-level productivities z_{tf}. We can no longer assume a log-normal productivity distribution, since, 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 the productivity distribution to be more log-convex than Pareto to fit the data. To achieve this, we assume that productivity is determined by a latent factor $\tilde{z}_{tf}$ that follows an AR(1) process as before and determines productivity as:

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

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

[24]The full specification for H and the derivation of equation (7) are provided in Appendix A.2.

$\tilde{z}_{tf}$ is a standard normal. This implies that $\ln z_{tf}$ follows a Weibull distribution with CDF

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

The Weibull distribution generalizes the exponential distribution by introducing the shape parameter n, which controls the log-convexity of the distribution. When $n = 1$, 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 distribution for z_{tf} is more log-convex than Pareto. As shown in Proposition 5, n is the key parameter determining the net effect of rising competition on an industry labor share.

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_ℓ	Comparative advantage	0.30	Manufacturing labor share (BLS/BEA)	0.67	0.67
	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 (2021). 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.

We calibrate the demand parameters $\{\sigma, \nu\}$ and the firm productivity distribution $\{\zeta, n\}$ to match the average manufacturing markup, the ratio of the (unweighted) mean firm labor share to the aggregate labor share, and the share of sales among the top 4 and top 20 firms in manufacturing in 1982. These parameters are jointly calibrated with the fixed cost of operation and its dispersion as well as with mean entrant productivity to match the same moments from Lee and Mukoyama (2015) used above, and after setting $\rho_z = 0.95$.[25] Panel

[25] As before, entrants draw a latent factor $\tilde{z}_{tf}$ from a normal distribution that differs from the long-run distribution of $\tilde{z}$ insofar as its mean is shifted to the left by $\ln \mu_e < 0$, calibrated to match the relative size of

II in Table 3 summarizes the parameters and moments informing their calibration.

We calibrate $\sigma = 6.0$ to match an aggregate markup of 1.15 and a super-elasticity ν/σ of 0.22 to match a 1.1 ratio between the unweighted average labor share of manufacturing firms and the aggregate manufacturing labor share before 1982 from Kehrig and Vincent (2021). Our calibration attributes all the differences in labor shares by firm size in the 1982 steady state to markups. The log-concave demand system implies that smaller firms have lower markups; therefore, the unweighted mean labor share across manufacturing firms will exceed the aggregate labor share of the sector by a factor that increases in the super-elasticity of demand. A super-elasticity of 0.22 matches the ratio of 1.1 in the data.[26]

For the productivity distribution, we calibrate $\{\zeta, n\}$ to match sales shares of the top 4 and the top 20 firms within manufacturing industries in 1982, which correspond to the top 1.1% and top 5.5% of firms in the sales distribution. Intuitively, a higher sales share among the top 5.5% firms indicates a thicker tail of the productivity distribution—a higher ζ. Conditional on the top 5.5% share, a higher sales share among the top 1.1% firms 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. We find that $n = 0.74$—a small deviation from Pareto and thus a moderate degree of log-convexity—fits the manufacturing sales concentration in 1982.

Retail: We follow the same calibration approach for retail, reported in Table 4. Relative to manufacturing, the main difference 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% firms) of 15% in 1982 relative to the (also high) share of the top 20 firms (the top 0.12% firms) 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 obtained from our production function estimates in Section 1.[27]

3.3 Quantifying the role of competition and capital–labor substitution

As before, we treat the economy in 1982 as being in steady state. We then calibrate the q and λ shocks required to match (i) the observed decline in the sectoral labor share, (ii) the observed decline in average capital prices by sector, and (iii) the increase in sales concentration. In addition, and as we did in the previous section, we calibrate the fixed cost

entrants.

[26] Our calibrated demand super-elasticity is close to the preferred estimate in Edmond, Midrigan and Xu (2022) of 0.16, who estimate it to match labor share dispersion by firm size in US Census data. Appendix A.4 provides a robustness exercise with a lower super-elasticity of 0.16; the quantitative conclusions are similar.

[27] For retail and other non-manufacturing sectors, we lack data on exit and entry rates, as well as the relative size of entrants and exiters. Thus, we target the same moments as in our manufacturing calibration. We also keep the demand super-elasticity of 0.22 and provide robustness checks in Appendix A.4.

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_ℓ	Comparative advantage	0.22	Retail labor share (BLS/BEA)	0.72	0.72
	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.

of automation per task to match (iv) the higher adoption rate of automation technologies among large firms documented in Acemoglu et al. (2022).

The key assumption is that there are no other forces affecting sales concentration or the labor share in manufacturing and retail. Relative to the previous section, we now account for the possibility that part of the decline in the labor share is due to rising competition (and part of the increase in sales concentration is due to lower capital prices), and so we now put the decline in capital prices and the increase in competition on equal footing. Our exercise answers 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? We conduct this exercise separately for manufacturing and for retail to illustrate that the answer to this question varies by sector.[28]

Manufacturing: Column 1 in Table 5 summarizes the manufacturing data and Column 2

[28] Other forces that potentially explain the declining labor share decline and rising concentration include changes in monopsony power and higher TFP dispersion. Berger, Herkenhoff and Mongey (2022) estimate that labor market concentration has decreased over time, raising the labor share by 3 pp since 1982. Likewise, the increase in concentration could reflect higher TFP dispersion (as in Akcigit and Ates, 2023; Olmstead-Rumsey, 2019). In our model, such forces increase concentration but barely affect the labor share. E.g., if the observed increase in sales concentration in manufacturing was due to rising TFP dispersion (modeled via an increase in ζ), we find a reduction in the sectoral labor share of only 1.4 pp. To match the manufacturing labor share decline one also needs top firms to become more capital-intensive over time, as in our theory.

reports our results. To match (i)–(iv), our model requires a uniform decline in the effective price of capital of $d\ln q_u = 0.6$, a skewed shock of $d\ln q_s = 4.57$, and a mild increase in market access of $d\ln\lambda = 0.06$. The reason why we calibrate a small increase in λ is that the rise in sales concentration in manufacturing 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). The 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 skewed q shock to generate the large decline in the manufacturing labor share.

Table 5: Labor share dynamics in manufacturing using the non-CES demand system (1982–2012)

		Model		
	Data	Baseline non-CES model	Only effects of $d\ln q$	Only effects of $d\ln\lambda$
	(1)	(2)	(3)	(4)
I. Parameters and inferred aggregate shocks				
$d\ln q_u$		0.60	0.60	0
$d\ln q_s$		4.57	4.57	0
$d\ln\lambda$		0.06	0	0.06
c_a		0.16	0.16	0.16
II. Targeted moments, 1982–2012				
Δ aggregate labor share	-0.199	-0.199	-0.209	0.004
Δ log average capital price	-1.081	-1.074	-0.994	0
Δ log 4 firms' sales share	0.140	0.143	0.078	0.070
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1.62	9.74
III. Typical firm labor share and other moments				
Δ median labor share	0.030	0.036	0.016	0.013
Δ unweighted mean	-0.017	-0.005	-0.025	0.012
Δ log 20 firms' sales share	0.072	0.142	0.105	0.046
Δ log productivity dispersion (from 1980-2000)	0.050	0.073	0.060	0.000
IV. Melitz–Polanec decomposition from Autor et al. (2020)				
Δ aggregate labor share	-0.185	-0.195	-0.205	0.003
Δ unweighted incumbent mean	-0.002	0.017	-0.001	0.014
Exit	-0.055	-0.015	-0.016	-0.018
Entry	0.059	0.012	0.014	0.015
Covariance term	-0.187	-0.210	-0.203	-0.008
V. Covariance decomposition from Kehrig and Vincent (2021)				
Market share dynamics	0.047	0.075	0.075	0.070
Labor share by size dynamics	-0.043	-0.056	-0.072	0.082
Cross-cross dynamics	-0.232	-0.206	-0.198	-0.147
VI. Markups, 1982–2012				
Δ log aggregate markup	-0.011	0.014	0.014	0.001
Within-firm change in markup	-0.076	-0.031	-0.025	-0.014
Reallocation to high-markup firms	0.065	0.045	0.039	0.015

Notes: Column (2) reports the findings from our benchmark model. Column (3) shows results when shutting down the market access shock, column (4) when shutting down instead the price of capital shocks.

In response to these shocks, our model provides a good fit to the manufacturing data, matching the aggregate labor share decline, the rise in concentration, and the untargeted behavior of the labor share behavior across firms.

A small fixed cost of automating tasks $c_a = 0.16$, calibrated to match the ABS adoption data, ensures that the labor share of the typical manufacturing firm increases despite lower capital prices. Panel III shows that the median firm labor share rises by 3.6 pp and the unweighted mean decreases by 0.5 pp during 1982–2012, which match the 3 pp increase in the median and the 1.7 pp decline in the unweighted mean in the data.

Panel IV of Table 5 describes the behavior of the labor share across firms using the Melitz–Polanec decomposition from Autor et al. (2020). As before, 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 covariance term contributed 18.7 pp to the labor share decline in the data and 21 pp in our model. Panel V shows that our model with endogenous markups quantitatively matches the covariance decomposition in Kehrig and Vincent (2021). The labor share decline is generated by firms that expand at the same time they reduce their labor shares (the cross-cross dynamics term), and not by the subsequent expansion of firms that had low labor shares initially (the market share effect). The cross-cross dynamics term contributed 23.2 pp to the labor share decline in the data and 20.6 pp in our model. This term is more negative than in the CES model because firms that expand not only automate more tasks when they do so, but also raise their markups during periods of expansion—a feature of Marshall's second law. As in the data, our model now produces a positive market share effect of 7.5 pp. This is because firms that had low labor shares in 1982 were already large. Due to mean reversion in the productivity process, these firms experience a decline in market share over time, contributing to an increase in the labor share.

The model also generates a transient behavior of firms' labor shares. Kehrig and Vincent (2021) document that firms experience cycles during which their labor shares decline and partly recover, and that these cycles have become more pronounced over time. In our model, these cycles are a natural implication of Marshall's second law. In the initial steady, firms that receive a positive productivity shock raise markups and lower their labor shares. These firms will then lower their markups and raise their labor shares as their productivity mean-reverts over time. Along the transition, these cycles become more pronounced and asymmetric, since firms also automate some of their tasks upon receiving a positive productivity shock. Figure 5 illustrates the phenomenon by plotting the impulse-response of firms' labor shares to a productivity shock in the initial steady state and along the transition.[29]

[29] Using data for a subset of industries for which price and quantity data are available (chemicals, aluminum

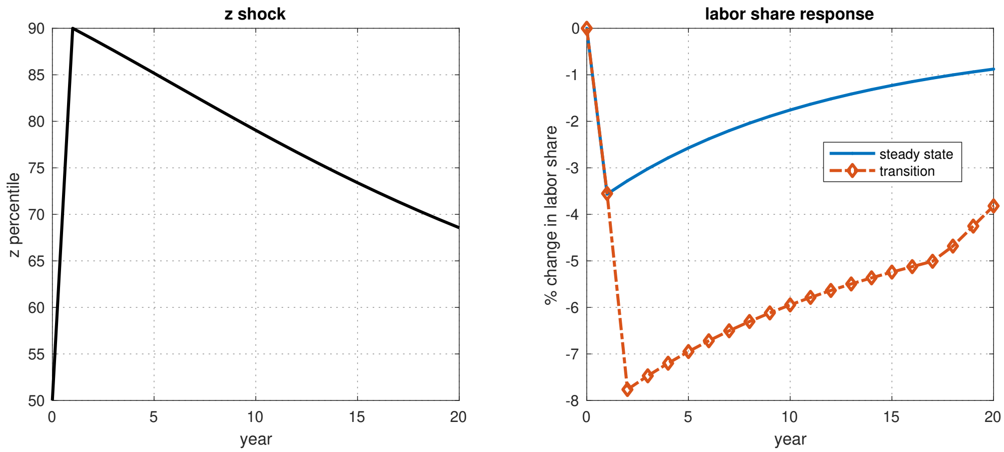

FIGURE 5: IMPULSE RESPONSE OF FIRM LABOR SHARE TO PRODUCTIVITY SHOCK. Left panel displays the shock: a firm's productivity z_{tf} jumps from the 50th to the 90th percentile on impact, and then mean-reverts according to ρ_z. Right panel displays firms' labor share response (relative to firm with median productivity).

The exercises in this section show that Marshall's second law of demand is an important ingredient needed to quantitatively match the joint dynamics of the labor share and sales shares documented in Kehrig and Vincent (2021). They also shows that our model can generate these firm dynamics as a response to lower capital prices.

To understand the contribution of the q and λ shocks, we provide counterfactual scenarios where we shut them down sequentially. In Column 3, we consider the effect of the q shocks holding λ constant at its initial steady state. The decline in effective capital prices generates all of the observed decline in the manufacturing labor share, and 7.8 log points of the observed 14 log points increase in sales concentration among the top 4 firms in the sector. The q shock alone generates all the firm-level patterns in panels IV and V. In Column 4, we consider the effect of the λ shock holding q constant at its initial steady state. The increase in competition does not contribute to the decline in the manufacturing labor share. Its role is to increase sales concentration among the top 4 firms by 7 log points.

Retail: We now calibrate the decline in capital prices and increase in market access required to match the labor share decline and the rising sales concentration in retail. Table 6 summarizes the retail data and presents the calibrated q (capital prices) and λ (market access) shocks. As before, we calibrate the fixed cost of automating tasks to match the higher adoption of automation technologies by large firms in the data.[30] We also require

sheets, and others), Kehrig and Vincent (2021) argue that firm prices increase during these cycles. One could capture this feature by adding firm-level demand shocks to our model. It is also important to note that these cycles occur naturally in a steady state (as shown in Figure 5), and so their mere presence does not discriminate between explanations of the aggregate labor share decline over time.

[30] Acemoglu et al. (2022) report similar relative adoption rates by size inside and outside of manufacturing.

that the q shocks generate an average decrease in capital prices of 86.5 log points, which is our quality-adjusted estimate for the decline in capital prices in this sector.

TABLE 6: Labor share dynamics in retail 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				
$d\ln q_u$		0.48	0.48	0
$d\ln q_s$		2.77	2.77	0
$d\ln \lambda$		0.30	0	0.30
c_a		0.06	0.06	0.06
II. Targeted moments, 1982–2012				
Δ aggregate labor share	-0.127	-0.127	-0.122	-0.022
Δ log average capital price	-0.865	-0.864	-0.690	0.000
Δ log sales concentration	0.546	0.546	0.065	0.480
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1.12	2.57
III. Typical firm labor share and other moments				
Δ median labor share		0.048	-0.036	0.037
Δ unweighted mean		0.028	-0.046	0.035
Δ log productivity dispersion		0.033	0.016	0.001
IV. Markups, 1982–2012				
Δ log aggregate markup	0.038	0.051	0.013	0.038
Within-firm change in markup	-0.026	-0.015	-0.010	-0.016
Reallocation to high-markup firms	0.064	0.066	0.023	0.054

Notes: Column (2) reports the findings from our benchmark model. Column (3) shows results when shutting down the competition shock, and column (4) when shutting down instead the price of capital shock.

In retail, we observe a labor share decline of 12.7 pp, a rise in sales 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 54.6 log points. For comparison, sales concentration increased by 14 log points in manufacturing. Our model generates the observed patterns for retail with a smaller uniform decline in the price of capital of $d\ln q_u = 0.48$, a skewed q shock of $d\ln q_s = 2.77$, and an increase in market access of $d\ln \lambda = 0.30$—an order of magnitude larger than in manufacturing.[31]

The large increase in market access calibrated for retail aligns with other studies. For example, Hsieh and Rossi-Hansberg (2023) document that retail firms increased their number of establishments by 21.6% and locations by 18.6% during 1977–2013.[32] Aghion et al. (2023)

[31] In retail and other non-manufacturing sectors, 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 small fraction (0.023%) of all retailers, whereas they account for 1.1% of all manufacturing firms.

[32] For manufacturing firms, the number of establishments and locations increased by less than 2%, which provides some support for the weaker λ shock estimated in that sector, although the tradable nature of manufacturing goods makes this less relevant for understanding changes in market access.

provide a model where lower overhead costs increase the number of markets serviced by top firms in retail, wholesale, and services by 31.3% since 1990.

The smaller skewed q shock in retail suggest that, in this sector, a larger share of advances in capital goods took place at the intensive margin, without leading to as much automation (and productivity dispersion) as in manufacturing.

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 market access to match the observed rise in 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 has a more pronounced effect on the labor share, leaving a smaller role for lower capital prices and the skewed q shock.

Our model provides a good fit to the available data for retail. As before, a small fixed cost of automating tasks matches the adoption data and leads to a rise in the labor share of the typical retail firm. We find that the unweighted mean labor share rises by 2.8 pp and the median by 4.8 pp. Though we lack direct data on these moments, this is in line with the evidence in Autor et al. (2020), who find that the unweighted mean of payroll shares in retail increased by 4.4 pp for 1982–2012.

Columns 3 and 4 report the effects of the increase in q and λ shocks separately. The large increase in market access in retail generates 20% of the decline in the labor share (−2.2 pp) and 90% of the increase in concentration. Increased market access also contributed to a rising labor share for the median and the unweighted mean of firms.

Other sectors: We also conducted our decomposition for other economic sectors, including wholesale, and utilities & transportation. The results are available in Appendix A.5. Relative to manufacturing, in wholesale and utilities & transportation, we infer a larger λ shock—required to match the large increase in concentration in these sectors—and a smaller skewed q shock—required to match the less pronounced labor share decline. The increase in market access plays a small role for the aggregate labor share of these sectors, since their calibrated productivity distribution is close to Pareto ($n = 0.74$ and $n = 0.75$).

3.4 Implications for Markups

This sub-section summarizes the predictions of our model for markups and compares them with the data. 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:

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

Panel VI of Table 5 summarizes the predictions of our model for markups in manufacturing. Our model predicts a mild increase in the manufacturing markup of 1.4% (from 1.150 to 1.166), which implies that markups account for 0.9 pp of the 20 pp decline in the manufacturing labor share. In line with Proposition 5, 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 3.1% 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 4.5% during this period. The results in columns (3) and (4) show that both lower capital prices and rising competition reallocate economic activity towards firms with high markups. The fact that rising competition leads to this form of reallocation is in line with Proposition 4. Lower capital prices have a similar reallocation effect because automation favors the expansion of large firms.

Panel IV of Table 6 summarizes the predictions of our model for markups in retail. Our model predicts an increase in the aggregate retail markup of 5.1% (from 1.15 to 1.21). This is the result of a 1.5% decrease in the within-firm component and a 6.6% increase driven by reallocation to high markup firms. Most of the rise in markups is explained by the sizable λ shock in this sector. In sum, markups explain 3.6 pp of the 12.7 pp decline in the labor share of retail, with the reallocation component generating 4.6 pp (36%) of the decline.

We now compare these predictions to estimates of markups from Compustat for 1980–2012. As is the case with any paper that relies on Compustat to obtain markups, these estimates must be interpreted with caution since the sample of firms is selected. We use the production function estimates from Section 1 to compute markups for firm f at time t as $\mu_{tf} = \varepsilon^v_{tf}/s^v_{tf}$, where s^v_{tf} denotes the share of variable inputs in revenue and ε^v_{tf} denotes the variable input elasticity (one minus the output–capital elasticity in Section 1).

Figure 6 plots the cumulative percent change in the inverse markup $1/\mu_t$ for manufacturing industries in the left panel and non-manufacturing industries in the right panel, decomposed into within-firm changes and the reallocation component.[33]

In line with our quantitative results, our estimates from Compustat support the conclusion that markups played no role on net in driving the decline of the manufacturing labor share. In this sector, the reallocation component accounts for a decline in the labor share of 6.5% between 1980 and 2012 (similar to the 4.5% decline in the labor share due to the reallocation towards large and high markup firms in our model). However, the estimated within-firm changes in markups go in the opposite direction and offset most of the reallocation component. This is in line with our calibration, where within-firm changes in markups and the reallocation component roughly cancel each other out in manufacturing (a consequence of the close to Pareto productivity distribution in this sector).[34]

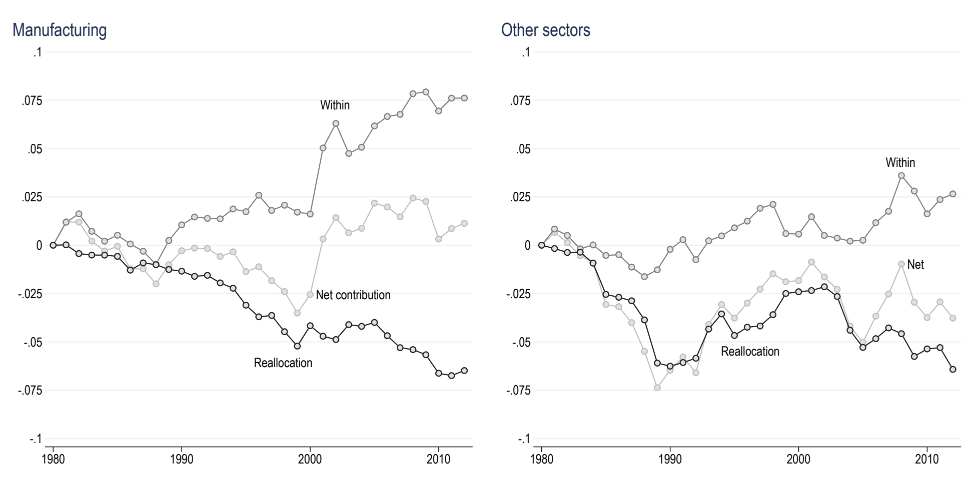

FIGURE 6: CONTRIBUTION OF WITHIN-FIRM CHANGES IN MARKUPS AND BETWEEN-FIRM REALLOCATION TO (PERCENT) CHANGES IN THE LABOR SHARE. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other sectors.

[33] 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 $-\ln \mu_t + \sum_f \omega_f \Delta \ln \mu_f$ and report their cumulative contributions over time.

[34] Figure S7 in the Appendix explains why we find a small role for markups relative to previous papers. First, we compute markups as an harmonic sales weighted mean across firms, which is the relevant weight to quantify their contribution to the labor share decline. Second, our production function estimates allow for differences in the factor intensity of technology across firms in the same industry. Even though our exercise relies on Compustat data, our conclusions receive support from recent work by Foster, Haltiwanger and Tuttle (2022) using establishment-level data from the Annual Survey of Manufactures for 1972-2014.

Outside manufacturing, the reallocation component reduced the labor share by 6.4% during 1980–2012, which accounts for a third of the observed decline in retail. This is in line with our quantitative model, where we estimate a 6.6% reduction in the retail labor share due to the reallocation of sales towards high markup firms, most of it in response to the λ shock. Moreover, in these sectors, the rise in the within component has been weak and only increased the labor share by 2.6% (similar to the 1.5% in our model). These estimates support the idea that, outside manufacturing, rising competition due to increased market access reduced the labor share via reallocation towards high markup firms.

4 Concluding remarks

This paper offered a quantitative exploration of the forces driving the decline in the labor share and the rise in concentration across US sectors. Our exercises show that it is possible to have a coherent and quantitative description of the manufacturing labor share decline driven by lower capital prices that fits the relevant micro and macro facts:

1. Our model generates the large observed decline in the manufacturing and retail labor share in response to q shocks that are of a size consistent with data on capital prices.

2. Our model generates these declines even in scenarios where an econometrician would estimate an elasticity of substitution between capital and labor below 1. Our explanation is therefore consistent with the limited substitution of capital for labor in response to higher wages estimated by Oberfield and Raval (2021) and others.

3. The model reproduces the striking fact that while the manufacturing labor share declined by 20 pp, the labor share of a typical firm increased or remained unchanged.

4. In response to q shocks, the model generates the firm-level patterns in Autor et al. (2020) and Kehrig and Vincent (2021) for manufacturing. We conclude that the new firm-level facts documented in Autor et al. (2020) and Kehrig and Vincent (2021) cannot rule out explanations of the labor share decline driven by lower capital prices.

5. Our model can also account for divergent sectoral experiences through different combinations of shocks. This exercise points to more sizable skewed q shocks in manufacturing and a bigger increase in market access in retail, wholesale, and utilities.

6. Except for retail, our model suggest that increased market access and competition had small effects on the aggregate sector labor share. Still, λ shocks are an important force behind the rise in sales concentration, especially outside manufacturing.

References

Acemoglu, D., G. W. Anderson, D. N. Beede, C. Buffington, E. E. Childress, E. Dinlersoz, L. S. Foster, N. Goldschlag, J. C. Haltiwanger, Z. Kroff, P. Restrepo, and N. Zolas (2022): *Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey* U. Chicago Press.

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., A. Manera, and P. Restrepo (2020): "Does the US Tax Code Favor Automation?," *Brookings Papers on Economic Activity*, 2020(1), 231–300.

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.

Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2023): "A Theory of Falling Growth and Rising Rents," *The Review of Economic Studies*.

Aghion, P., B. F. Jones, and C. I. Jones (2018): *Artificial Intelligence and Economic Growth*pp. 237–282. University of Chicago Press.

Akcigit, U., and S. Ates (2023): "What Happened to U.S. Business Dynamism?," *Journal of Political Economy*, (forthcoming).

Akerman, A., E. Leuven, and M. Mogstad (2022): "Information Frictions, Internet, and the Relationship between Distance and Trade," *American Economic Journal: Applied Economics*, 14(1), 133–63.

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 (2022): "Labor Market Power," *American Economic Review*, 112(4), 1147–93.

Bonfiglioli, A., R. Crinò, H. Fadinger, and G. Gancia (2020): "Robot Imports and Firm-Level Outcomes," Discussion paper, CEPR Discussion Papers No. 14593.

Brynjolfsson, E., C. Buffington, N. Goldschlag, J. F. Li, J. Miranda, and R. Seamans (2023): "The Characteristics and Geographic Distribution of Robot Hubs in U.S. Manufacturing Establishments," Working Paper 31062, National Bureau of Economic Research.

Cheng, H., L. A. Drozd, R. Giri, M. Taschereau-Dumouchel, and J. Xia (2021): "The Future of Labor: Automation and the Labor Share in the Second Machine Age," Discussion paper, FRB Philadelphia.

Clementi, G. L., and B. Palazzo (2016): "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," *American Economic Journal: Macroeconomics*, 8(3), 1–41.

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*, 110(12), 3952–90.

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.

Drozd, L., M. Taschereau-Dumouchel, and M. Mendes Tavares (2022): "Understanding Growth Through Automation: The Neoclassical Perspective," Discussion paper, FRB of Philadelphia.

Edmond, C., V. Midrigan, and D. Y. Xu (2022): "How Costly Are Markups?," *Journal of Political Economy*, forthcoming.

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.

Foster, L. S., J. C. Haltiwanger, and C. Tuttle (2022): "Rising Markups or Changing Technology?," Working Paper 30491, National Bureau of Economic Research.

Gordon, R. J. (1990): *The Measurement of Durable Goods Prices*. University of Chicago Press.

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

Hopenhayn, H., J. Neira, and R. Singhania (2022): "From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," *Econometrica*, 90(4), 1879–1914.

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(5), 1127–1150.

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

Hsieh, C.-T., and E. Rossi-Hansberg (2023): "The Industrial Revolution in Services," *Journal of Political Economy Macroeconomics*, 1(1), 3–42.

Hubmer, J. (2023): "The Race Between Preferences and Technology," *Econometrica*, 91(1), 227–261.

Humlum, A. (2019): "Robot Adoption and Labor Market Dynamics," Discussion paper, MIMEO, Princeton University.

Jones, B. F., and X. Liu (2022): "A Framework for Economic Growth with Capital-Embodied Technical Change," Working Paper 30459, National Bureau of Economic Research.

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

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 (2021): "The Micro-Level Anatomy of the Labor Share Decline," *The Quarterly Journal of Economics*, 136(2), 1031–1087.

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 (2021): "Robots and Firms," *The Economic Journal*, 131(638), 2553–2584.

Kogan, L., D. Papanikolaou, L. D. W. Schmidt, and B. Seegmiller (2021): "Technology-Skill Complementarity and Labor Displacement: Evidence from Linking Two Centuries of Patents with Occupations," Working Paper 29552, National Bureau of Economic Research.

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.

Oberfield, E., and D. Raval (2021): "Micro data and macro technology," *Econometrica*, 89(2), 703–732.

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.

Peters, M., and C. Walsh (2021): "Population Growth and Firm Dynamics," Working Paper 29424, National Bureau of Economic Research.

Philippon, T. (2019): *The Great Reversal: How America Gave Up on Free Markets*. Harvard University.

Zeira, J. (1998): "Workers, Machines, and Economic Growth," *The Quarterly Journal of Economics*, 113(4), 1091–1117.

Online Appendix to "Capital–Labor Substitution and Firms' Labor Shares"

Joachim Hubmer and Pascual Restrepo

A.1 Technical proofs for the CES-demand model

This section provides proofs for Propositions 1, 2, and 3. In addition, we provide an additional lemma characterizing the stationary equilibrium of the economy.

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

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

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

$$\Omega_t(\alpha, z) = \pi_t'(\alpha) \cdot \mathbb{E}[z'^{\sigma-1}|z] + \mathbb{E}\left[P_t(z') \cdot \min\left\{c_a \cdot y_t, \frac{1}{1+r}\Omega_{t+1}(\alpha, z')\right\}\Big|z\right], \tag{A1}$$

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 $n \leq N$. 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, $P_t(z') \cdot \min\{c_a \cdot$

$y_t, (1/(1+r)) \cdot \Omega^{(N)}_{t+1}(\alpha, z')\}$ is (weakly) increasing in z' (due to the inductive hypothesis), and so the term $\mathbb{E}\Big[P_t(z') \cdot \min\{c_a \cdot y_t, (1/(1+r)) \cdot \Omega^{(N)}_{t+1}(\alpha, z')\}|z\Big]$ also (weakly) increases in z, which completes the inductive step.

Because the set of weakly increasing functions is closed, $\Omega_t(\alpha, z) = \lim_{n\to\infty} \Omega^{(n)}_t(\alpha, z)$ 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 (A1) 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 (A1) 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 effective capital productivity, $q_t(x) = q(x)$, 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

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

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

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

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

Lemma A1 *Given a constant level of effective capital productivity $q_t(x) = q(x)$, 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 A1 justifies our focus on steady states where all firms operate a technology $\alpha^*(w^*(q);q)$ and wages are given by $w^*(q)$. Propositions 2 and 3 explore how factor shares vary across these steady states in response to different changes in capital prices.

Proof of Proposition 2. 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

$$\tilde{c}(w, q, \alpha^*) = \left(\Gamma_0^k(\alpha^*) \cdot q^{\eta-1} + \Gamma^\ell(\alpha^*) \cdot w^{1-\eta}\right)^{\frac{1}{1-\eta}}$$

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

$$\varepsilon^\ell = \frac{\Gamma^\ell(\alpha^*) \cdot w^{1-\eta}}{\Gamma_0^k(\alpha^*) \cdot q^{\eta-1} + \Gamma^\ell(\alpha^*) \cdot w^{1-\eta}}.$$

This common cost share for labor will vary with prices and α^*. Equation (2) implies that the change in the optimal threshold α^* satisfies

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

Using this expression for $d \ln \alpha^*$ and the definition of η^*, we obtain

$$\begin{aligned} d \ln \varepsilon^\ell &= \varepsilon^k \cdot d \ln \frac{\varepsilon^\ell}{\varepsilon^k} \\ &= \varepsilon^k \cdot (1-\eta) \cdot (d \ln q + d \ln w) + \varepsilon^k \cdot \frac{\partial \ln \Gamma^\ell(\alpha^*)/\Gamma_0^k(\alpha^*)}{\partial \ln \alpha} \cdot d \ln \alpha^* \\ &= \varepsilon^k \cdot (1-\eta) \cdot (d \ln q + d \ln w) + \varepsilon^k \cdot (\eta - \eta^*) \cdot (d \ln q + d \ln w) \\ &= \varepsilon^k \cdot (1-\eta^*) \cdot (d \ln q + d \ln w), \end{aligned}$$

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. We have that

$$\begin{aligned} d \ln \varepsilon^\ell_{tf} &= \varepsilon^k \cdot d \ln \frac{\varepsilon^\ell}{\varepsilon^k} \\ &= \varepsilon^k \cdot (1-\eta) \cdot (d \ln q + d \ln w_t) + \varepsilon^k \cdot \frac{\partial \ln \Gamma^\ell(\alpha^*)/\Gamma_0^k(\alpha^*)}{\partial \ln \alpha} \cdot d \ln \alpha_{tf} \\ &= \varepsilon^k \cdot (1-\eta) \cdot (d \ln q + d \ln w_t) + \varepsilon^k \cdot \frac{\partial \ln \Gamma^\ell(\alpha^*)/\Gamma_0^k(\alpha^*)}{\partial \ln \alpha} \cdot d \ln \alpha^* \cdot \frac{d \ln \alpha_{tf}}{d \ln \alpha^*} \\ &= \varepsilon^k \cdot (1-\eta) \cdot (d \ln q + d \ln w_t) + \varepsilon^k \cdot (\eta - \eta^*) \cdot (d \ln q + d \ln w_t) \cdot \frac{d \ln \alpha_{tf}}{d \ln \alpha^*} \\ &= \varepsilon^k \cdot \left(1 - \eta + (\eta^* - \eta) \frac{d \ln \alpha_{tf}}{d \ln \alpha^*}\right) \cdot (d \ln q + d \ln w_t), \end{aligned}$$

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 3. Let $q(x) = q \cdot q_0(x)$ for $x > \alpha^*$ and $q(x) = q_0(x)$ otherwise. We are interested in comparative statics as q changes from 1 by $d \ln q$.

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

We now turn cost shares. In steady state, all firms have the same labor cost share

$$\varepsilon^\ell = \frac{\Gamma^\ell(\alpha^*) \cdot w^{1-\eta}}{\Gamma_0^k(\alpha^*) + \Gamma^\ell(\alpha^*) \cdot w^{1-\eta}}.$$

This common cost share for labor will vary with prices and α^*. Equation (2) implies that the change in the optimal threshold α^* satisfies

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

Using this expression for $d\ln\alpha^*$ and the definition of η^*, we obtain

$$\begin{aligned}
d\ln\varepsilon^\ell &= \varepsilon^k \cdot d\ln\frac{\varepsilon^\ell}{\varepsilon^k} \\
&= \varepsilon^k \cdot (1-\eta)\cdot d\ln w + \varepsilon^k \cdot \frac{\partial \ln \Gamma^\ell(\alpha^*)/\Gamma_0^k(\alpha^*)}{\partial \ln \alpha} \cdot d\ln\alpha^* \\
&= \varepsilon^k \cdot (1-\eta)\cdot d\ln w - \varepsilon^k \cdot (\eta^* - \eta)\cdot (d\ln q + d\ln w).
\end{aligned}$$

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. We have that

$$\begin{aligned}
d\ln\varepsilon^\ell_{tf} &= \varepsilon^k \cdot d\ln\frac{\varepsilon^\ell}{\varepsilon^k} \\
&= \varepsilon^k \cdot (1-\eta)\cdot d\ln w_t + \varepsilon^k \cdot \frac{\partial \ln \Gamma^\ell(\alpha^*)/\Gamma_0^k(\alpha^*)}{\partial \ln \alpha} \cdot d\ln\alpha_{tf} \\
&= \varepsilon^k \cdot (1-\eta)\cdot d\ln w_t + \varepsilon^k \cdot \frac{\partial \ln \Gamma^\ell(\alpha^*)/\Gamma_0^k(\alpha^*)}{\partial \ln \alpha} \cdot d\ln\alpha^* \cdot \frac{d\ln\alpha_{tf}}{d\ln\alpha^*} \\
&= \varepsilon^k \cdot (1-\eta)\cdot d\ln w_t - \varepsilon^k \cdot (\eta^* - \eta)\cdot (d\ln q + d\ln w)\cdot \frac{d\ln\alpha_{tf}}{d\ln\alpha^*}.
\end{aligned}$$

■

A.2 Technical proofs for the model with variable markups

This section provides the proofs of the theoretical results in section 3.

A.2.1 Implications of Marshall's weak and strong second laws

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

Lemma A2 *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 A2. 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 $p^*(c)$ increases 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

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

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 $\ln c$. Moreover, $\ln \mu^*(c) = \ln p^*(c) - \ln c$ inherits 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. ■

A.2.2 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.

First, we show that the demand system that results from aggregating customers' demand is equivalent to assuming a single representative customer that demands varieties to produce a final good y_t which can then be used for consumption or to produce capital goods. In particular, the customers introduced in the main text could also represent producers of capital goods, and that is why we do not refer to them simply as consumers. This representative customer solves 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

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

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

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

which pins down the competitors' price index ρ_t. These three equations coincide with equations (5), (4), and (6) and establish the aggregation result.

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 capital productivity $q_t(x)$, market access, 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 (6) holds.

E2. The competitors' price index condition in equation (5) 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

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

E5. Entry decisions maximize the value of entrants

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

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 entry and exit.

Proof of Proposition 4. 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 A2 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,

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

is increasing in $\bar{c}$ for $z > z'$. ■

Proof of Proposition 5. 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

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

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

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

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

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

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

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

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 $\bar{c}$. ■

A.2.3 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

$$-\frac{D'(\bar{p}_{tf}) \cdot \bar{p}_{tf}}{D(\bar{p}_{tf})} = \frac{\sigma}{1 - \nu \cdot \ln\left(\bar{p}_{tf} \cdot \frac{\sigma}{\sigma-1}\right)} = \sigma \cdot D(\bar{p}_{tf})^{-\frac{\nu}{\sigma}}, \tag{A2}$$

which reduces to the constant σ if $\nu = 0$ (the benchmark case of a CES aggregator). In general, equation (A2) 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 conclude, we show that the Klenow-Willis aggregator satisfies Marshall's second laws. Equation (7) 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.

A.3 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 η^*.

The induced elasticity of substitution is, as usual, defined implicitly by the identity

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

This gives the change in labor-to-capital cost ratios resulting from an increase in wages. Higher wages reduce the labor-to-capital cost ratio if $\eta^* > 1$ and increase it if $\eta^* < 1$.

Define $\alpha^*(w)$ implicitly as in the text by the solution to

$$w = \frac{\psi(\alpha^*(w))}{q(\alpha^*(w))}.$$

Define $h_k(w) = \Gamma^k(\alpha^*(w))$ and $h_\ell(w) = \Gamma^\ell(\alpha^*(w))$. Differentiating these yields

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

which combined yield the differential equation

$$h_\ell'(w) = -\,w^{\eta-1}\cdot h_k'(w). \tag{A3}$$

In our model, the ratio of labor to capital costs can be computed as

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

Hence, the induced elasticity of substitution is constant and equal to η^* if and only if

$$\frac{h_\ell(w)}{h_k(w)} = \chi\cdot w^{\eta-\eta^*}. \tag{A4}$$

This equation implies that $\eta^* > \eta$, since $h_\ell'(w) < 0$ and $h_k'(w) > 0$. Rearranging and taking derivatives yields

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

Combining this equation with (A3) yields the differential equation for $h_k(w)$:

$$\text{(A5)} \qquad \frac{h_k'(w)}{h_k(w)} = (\eta^* - \eta) \cdot \frac{\chi \cdot w^{-\eta^*}}{\chi \cdot w^{1-\eta^*} + 1}.$$

This differential equation has two solutions, one for $\eta^* = 1$ and another one for $\eta^* \neq 1$.

A.3.1 Parametrization with $\eta^* = 1$:

Our baseline calibration focuses on the case $\eta^* = 1$. In this case, we can write (A5) as

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

Integrating both sides we get

$$h_k(w) = M \cdot w^{(1-\eta) \cdot \frac{\chi}{\chi+1}},$$

for some $M > 0$, and

$$h_\ell(w) = M \cdot \chi \cdot w^{(1-\eta) \cdot \frac{-1}{\chi+1}}.$$

With the functions $h_k(w)$ and $h_\ell(w)$ 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 $\alpha^*(w) = f(w)$ from the positive reals to $[0, 1]$, capturing the optimal threshold rule, we can define

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

which also implies $\Gamma^k(x) = h_k(f^{-1}(x))$ and $\Gamma^\ell(x) = h_\ell(f^{-1}(x))$.

Taking the function $\alpha^*(w) = \left(\frac{w}{1+w}\right)^{\frac{1-\eta}{\gamma_\ell+\gamma_k}}$ yields the parametrization used in the text.

One can also verify that the implied share parameters are

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

the optimal automation decision satisfies

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

and the ratio of labor to capital costs is given by

$$\frac{s^{\ell}}{s^{k}} = \frac{\Gamma^{\ell}(\alpha^*)}{\Gamma^{k}(\alpha^*)} \cdot w^{1-\eta} = \frac{\gamma_k}{\gamma_\ell}.$$

This shows that, once firms are allowed to adjust their tasks, the labor share in costs is constant and equal to $\gamma_k/(\gamma_k + \gamma_\ell)$.

A.3.2 Parametrization with $\eta^* \neq 1$:

Integrating both sides of equation A5 gives the unique solution for $h_k(w)$

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

Using equation (A4), we also obtain

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

As before, we can now generate all possible parametrizations of the productivity schedules for capital and labor in $[0,1]$ that induce a constant elasticity of substitution $\eta^* \neq 1$ by picking an increasing function $\alpha^*(w) = f(w)$ from the positive reals to $[0,1]$.

For the exercise reported in Column 5 of Table 1, we take $\alpha^*(w) = h_k(w)$, which is an increasing function from the positive reals to $[0,1]$ when $\eta^* > 1$. This choice of $\alpha^*(w)$ implies $\phi(x) = 1$ and

$$\psi(x) = A \cdot \left(x^{\frac{1-\eta^*}{\eta^*-\eta}} - 1\right)^{\frac{1}{1-\eta^*}},$$

where A is a scaling factor that captures labor-augmenting technical change. 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

$$y_f = z_f \cdot \left(\alpha_f^{\frac{1}{\eta}} \cdot (q_0 \cdot k_f)^{\frac{\eta-1}{\eta}} + \Gamma^{\ell}(\alpha_f)^{\frac{1}{\eta}} \cdot (A \cdot \ell_f)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{1-\eta}},$$

$$\text{with } \Gamma^{\ell}(\alpha_f) = \left(1 - \alpha_f^{\frac{\eta^*-1}{\eta^*-\eta}}\right)^{\frac{\eta^*-\eta}{\eta^*-1}}. \tag{A6}$$

A.3.3 Modeling the q-shock:

For any parametrization of productivity schedules, define the combined q shock as

$$q_{I,t}(x) = \begin{cases} q_{u,t} & \text{if } x \leq \alpha_0^* \\ q_{u,t} \cdot \min\{\frac{\psi(x)/\psi(\alpha_0^*)}{\phi(x)/\phi(\alpha_0^*)}, q_{s,t}\} & \text{if } x > \alpha_0^* \end{cases}$$

for some increasing $\{q_{u,t}, q_{s,t}\}$ in $(1, \infty)$ converging to $\{q_u, q_s\}$.

In the main text, we also normalize baseline wages w_0 to 1 and assume that $q_0^I(x) = 1$, which implies $\psi(\alpha_0^*)/\phi(\alpha_0^*) = 1$, and simplifies the formulation of the shock.

Figure A1 represents the q shock graphically. Tasks are arranged in [0,1] in the horizontal axis. At time 0, we have $q_{I,t}(x) = 1$ for all tasks. Over time, the productivity with which the economy can produce the capital needed for task x rises, and the increase is more pronounced for more complex tasks. The term $\min\{\frac{\psi(x)/\psi(\alpha_0^*)}{\phi(x)/\phi(\alpha_0^*)}, q_{s,t}\}$ ensures that the shock approximates a series of small skewed q shocks like the ones considered in Proposition 3.

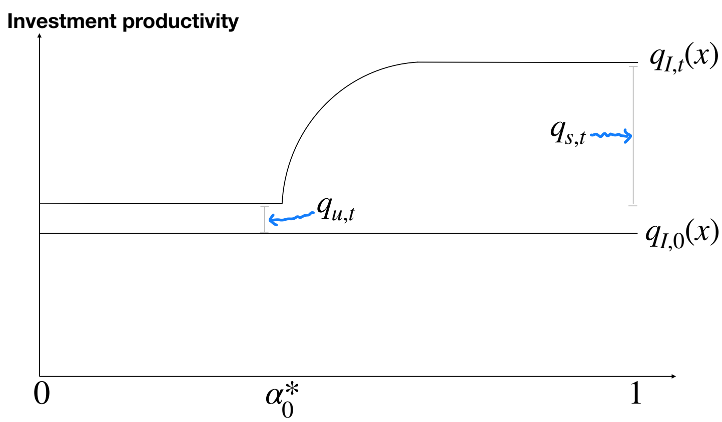

FIGURE A1: REPRESENTATION OF q SHOCKS. Tasks are arranged in [0,1] in the horizontal axis. At time 0, we have $q_{I,0}(x) = 1$ for all tasks. Over time, the productivity with which the economy can produce the capital needed for task x rises to $q_{I,t}(x)$.

Let α_0^* denote the optimal threshold at the initial steady state, α^* the new steady-state level after the combined q shock, and $\bar{\alpha}$ the point at which $\frac{\psi(x)/\psi(\alpha^*)}{\phi(x)/\phi(x^*)} = q_s$.

Here, α^* and $\bar{\alpha}$ are defined implicitly by

$$w \cdot q_u \cdot q_s = \frac{\psi(\alpha^*)}{\phi(\alpha^*)} \qquad\qquad w_0 \cdot q_s = \frac{\psi(\bar{\alpha})}{\phi(\bar{\alpha})}.$$

This implies that $\alpha_0^* < \bar{\alpha} < \alpha^*$, since wages increase with advances in technology.

After the shock, the task share for labor is $\Gamma_\ell(\alpha^*)$, which is lower than before, and the task share for capital is higher than before and given by

$$\begin{aligned}\Gamma_k(\alpha^*) =& q_u^{\eta-1} \cdot \Gamma_k(\alpha_0^*) + q_u^{\eta-1} \cdot w_0^{1-\eta} \cdot \int_{\alpha_0^*}^{\bar{\alpha}} \psi(x)^{\eta-1} \cdot dx + q_u^{\eta-1} \cdot q_s^{\eta-1} \cdot [\Gamma_k(\alpha^*) - \Gamma_k(\bar{\alpha})], \\ =& q_u^{\eta-1} \cdot \Gamma_k(\alpha_0^*) + q_u^{\eta-1} \cdot w_0^{1-\eta} \cdot [\Gamma_\ell(\alpha_0^*) - \Gamma_\ell(\bar{\alpha})] + q_u^{\eta-1} \cdot q_s^{\eta-1} \cdot [\Gamma_k(\alpha^*) - \Gamma_k(\bar{\alpha})].\end{aligned}$$

To conclude, we show that in our parametrization with $\eta^* = 1$ in the initial steady state, the combined q shock takes the economy to a new steady state with: i. a lower labor share and ii. an induced elasticity of substitution below 1.

To show (i), note that in the new steady state

$$\begin{aligned}\frac{\varepsilon_k}{\varepsilon_\ell} =& \frac{q_u^{\eta-1} \cdot \Gamma_k(\alpha_0^*) + q_u^{\eta-1} \cdot w_0^{1-\eta} \cdot [\Gamma_\ell(\alpha_0^*) - \Gamma_\ell(\bar{\alpha})] + q_u^{\eta-1} \cdot q_s^{\eta-1} \cdot [\Gamma_k(\alpha^*) - \Gamma_k(\bar{\alpha})]}{w^{1-\eta} \cdot \Gamma_\ell(\alpha^*)} \\ =& \frac{\Gamma_k(\alpha_0^*) + w_0^{1-\eta} \cdot [\Gamma_\ell(\alpha_0^*) - \Gamma_\ell(\bar{\alpha})] + q_s^{\eta-1} \cdot [\Gamma_k(\alpha^*) - \Gamma_k(\bar{\alpha})]}{q_u^{1-\eta} \cdot w^{1-\eta} \cdot \Gamma_\ell(\alpha^*)} \\ =& \frac{1 - w_0^{1-\eta} \cdot \Gamma_\ell(\bar{\alpha}) - q_s^{\eta-1} \cdot \Gamma_k(\bar{\alpha})}{q_u^{1-\eta} \cdot w^{1-\eta} \cdot \Gamma_\ell(\alpha^*)} + \frac{q_s^{\eta-1} \cdot \Gamma_k(\alpha^*)}{q_u^{1-\eta} \cdot w^{1-\eta} \cdot \Gamma_\ell(\alpha^*)}.\end{aligned}$$

By construction, the second term is constant and equal to the initial ratio $\frac{\varepsilon_k}{\varepsilon_\ell}$. The first term is positive, which shows that the labor share is lower in the new steady state.

For (ii), we show that $\frac{\varepsilon_k}{\varepsilon_\ell}$ decreases in w, which implies an induced elasticity below 1. It is enough to show that the first term in the expression for $\frac{\varepsilon_k}{\varepsilon_\ell}$ decreases in w. This term can be written as

$$\begin{aligned}&\frac{1 - w_0^{1-\eta} \cdot \Gamma_\ell(\bar{\alpha}) - q_s^{\eta-1} \cdot \Gamma_k(\bar{\alpha})}{q_u^{1-\eta} \cdot w^{1-\eta} \cdot \Gamma_\ell(\alpha^*)} = \\ &\quad \left(1 - w_0^{1-\eta} \cdot \Gamma_\ell(\bar{\alpha}) - q_s^{\eta-1} \cdot \Gamma_k(\bar{\alpha})\right) \cdot \frac{\Gamma_k(\alpha^*)}{q_u^{1-\eta} \cdot w^{1-\eta} \cdot \Gamma_\ell(\alpha^*)} \cdot \frac{1}{\Gamma_k(\alpha^*)}.\end{aligned}$$

The first term is a positive constant (since this is zero for the higher wage level that would result if all firms automated all tasks up to $\bar{\alpha}$), the second term is constant (by construction), and the third term decreases in w, as claimed.

A.4 The 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 (2022). 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 A1 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 $d \ln \lambda = 0.09$ (instead of $d \ln \lambda = 0.06$ 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 A1, we infer a rise in the aggregate markup of 1.0% in manufacturing (as opposed to 1.4% in Table 5) and of 2.6% in retail (as opposed to 5.1% in Table 6). The other main results, in particular the contribution of falling capital prices to the labor share decline, are similar across parameterizations.

Table A1: Model robustness: Labor share 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*			
$d \ln q_u$		0.63	0.63	0
$d \ln q_s$		4.66	4.66	0
$d \ln \lambda$		0.09	0	0.09
c_a		0.17	0.17	0.17
	II. Targeted moments, 1982–2012			
Δ aggregate labor share	-0.199	-0.199	-0.209	0.004
Δ log average capital price	-1.081	-1.074	-0.994	0.000
Δ log 4 firms' sales share	0.140	0.143	0.078	0.070
Relative adoption (P99+ vs. P50-75 firms)	1.708	1.711	1.621	9.744
	III. Typical firm labor share and other moments			
Δ median labor share	0.030	0.036	0.016	0.013
Δ unweighted mean	-0.017	-0.005	-0.025	0.012
Δ log 20 firms' sales share	0.072	0.142	0.105	0.046
Δ log productivity dispersion	0.050	0.073	0.060	0.000
	IV. Markups, 1982–2012			
Δ log aggregate markup	-0.011	0.010	0.011	0.000
Within-firm change in markup	-0.076	-0.025	-0.021	-0.010
Reallocation to high-markup firms	0.065	0.035	0.032	0.010
B. Retail (1982–2012)	*I. Parameters and inferred aggregate shocks*			
$d \ln q_u$		0.44	0.44	0
$d \ln q_s$		3.30	3.30	0
$d \ln \lambda$		0.48	0	0.48
c_a		0.08	0.08	0.08
	II. Targeted moments, 1982–2012			
Δ aggregate labor share	-0.127	-0.127	-0.142	-0.008
Δ log average capital price	-0.865	-0.869	-0.656	0.000
Δ log sales concentration	0.546	0.547	0.067	0.493
Relative adoption (P99+ vs. P50-75 firms)	1.708	1.71	1.29	2.68
	III. Typical firm labor share and other moments			
Δ median labor share		0.049	-0.030	0.045
Δ unweighted mean		0.029	-0.045	0.042
Δ log productivity dispersion		0.040	0.018	0.001
	IV. Markups, 1982–2012			
Δ log aggregate markup	0.038	0.026	0.010	0.019
Within-firm change in markup	-0.026	-0.029	-0.013	-0.020
Reallocation to high-markup firms	0.064	0.055	0.023	0.039

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 $d \ln q_s$, $d \ln q_u$, $d \ln \lambda$ and the automation fixed cost c_a over the transition.

A.5 Calibration of the non-CES demand model for other sectors

Table A2 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 A2: Steady state calibration of the non-CES demand model: Wholesale, Utilities & Transportation

	Parameter		Moment	Data	Model
	I. Wholesale: steady state parameters and moments (1982)				
γ_ℓ	Comparative advantage	0.63	Wholesale labor share	0.53	0.53
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)				
γ_ℓ	Comparative advantage	0.72	Util.&transp. labor share	0.51	0.51
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 A3 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 skewed capital price shocks ($d \ln q_s$) is small, while the inferred increase in competition ($d \ln \lambda$) is weaker than in retail but stronger than in manufacturing. The inferred automation fixed costs (c_a) are comparable.

TABLE A3: Labor share dynamics with non-CES demand: Wholesale, Utilities & Transportation

	(1) DATA	MODEL (2) BENCHMARK	(3) ONLY $d \ln q$	(4) ONLY $d \ln \lambda$
A. Wholesale (1982–2012)	*I. Parameters and inferred aggregate shocks*			
$d \ln q_u$		1.30	1.30	0
$d \ln q_s$		1.06	1.06	0
$d \ln \lambda$		0.19	0	0.19
c_a		0.18	0.18	0.18
	II. Targeted moments, 1982–2012			
Δ aggregate labor share	-0.045	-0.045	-0.048	0.005
Δ log average capital price	-1.596	-1.593	-1.563	0
Δ log sales concentration	0.202	0.209	0.048	0.163
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1.42	3.52
	III. Typical firm labor share and other moments			
Δ median labor share		0.187	0.156	0.033
Δ unweighted mean		0.150	0.122	0.031
Δ log productivity dispersion		0.182	0.165	0.102
	IV. Markups, 1982–2012			
Δ log aggregate markup	0.038	0.009	0.006	0.003
Within-firm change in markup	-0.026	-0.043	-0.032	-0.015
Reallocation to high-markup firms	0.064	0.052	0.039	0.018
B. Utilities & Transportation (1992–2012)	*I. Parameters and inferred aggregate shocks*			
$d \ln q_u$		0.62	0.62	0
$d \ln q_s$		0.44	0.44	0
$d \ln \lambda$		0.12	0	0.12
c_a		0.10	0.10	0.10
	II. Targeted moments, 1982–2012			
Δ aggregate labor share	-0.028	-0.028	-0.029	0.002
Δ log average capital price	-0.684	-0.683	-0.676	0
Δ log sales concentration	0.108	0.105	0.025	0.080
Relative adoption (P99+ vs. P50-75 firms)	1.71	1.71	1.47	4.53
	III. Typical firm labor share and other moments			
Δ median labor share		0.102	0.088	0.017
Δ unweighted mean		0.083	0.070	0.016
Δ log productivity dispersion		0.117	0.107	0.089
	IV. Markups, 1982–2012			
Δ log aggregate markup	0.038	0.004	0.002	0.001
Within-firm change in markup	-0.026	-0.015	-0.010	-0.008
Reallocation to high-markup firms	0.064	0.018	0.012	0.010

Notes: Column (2) contains the benchmark model. Due to data availability, the transition is over 1982–2012 for wholesale, resp. 1992–2012 for utilities & transportation. Column (3) shuts down the competition shock, and column (4) shuts down the price of capital shock.

Figure A2 and Table A4 summarize our findings and contrast them with our results for manufacturing and retail.

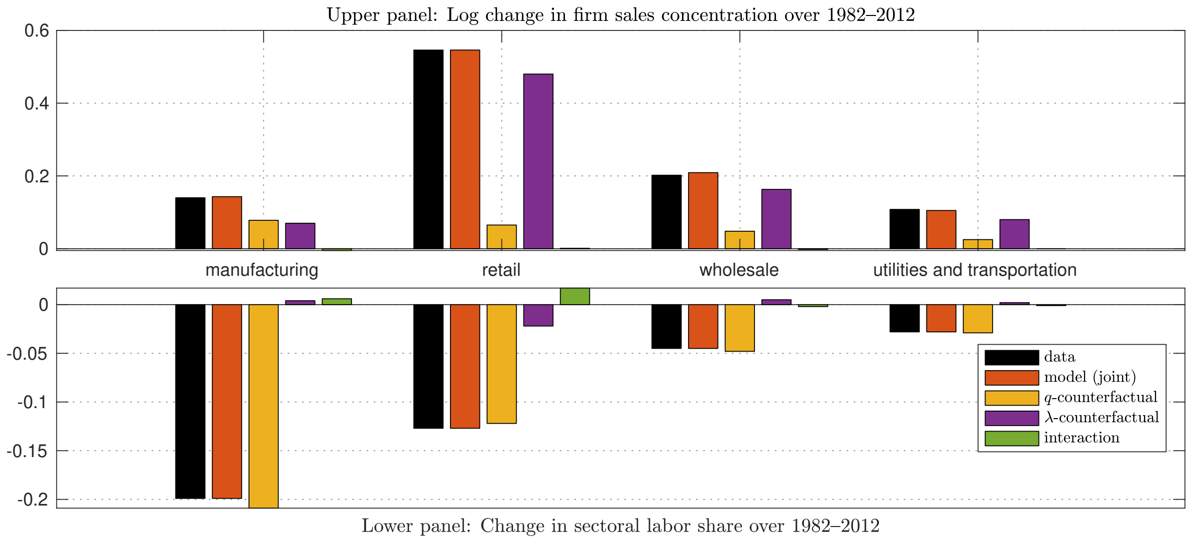

Figure A2: Model decomposition of labor share and sales concentration changes.

For each sector, the upper panel displays the log change in firm sales concentration (i) in the data (from Autor et al., 2020), (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 A3 for details.

Table A4: Inferred sectoral price of capital and market access shocks

	Decline in price of capital		Increase in market access
	$d\ln q_u$	$d\ln q_s$	$d\ln\lambda$
Manufacturing	0.60	4.57	0.06
Retail	0.48	2.77	0.30
Wholesale	1.30	1.06	0.19
Utilities & transportation	0.62	0.44	0.12

Notes: The table reports the inferred q and λ shocks required to match the sectoral patterns in manufacturing, retail, wholesale, and utilities & transportation.

Appendix References

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.

Edmond, C., V. Midrigan, and D. Y. Xu (2022): "How Costly Are Markups?," *Journal of Political Economy*, forthcoming.

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(5), 1127–1150.

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.

Supplementary Material to “Capital–Labor Substitution and Firms’ Labor Shares”

Joachim Hubmer and Pascual Restrepo

S.1 Estimating Output Elasticities

S.1.1 Data description, sample, and definitions

We use data from Compustat for 1960–2016. 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 follow De Loecker, Eeckhout and Unger (2020) and measure the capital stock using the gross value of property, plants, and equipment—*PPEGT* in Compustat. We obtained similar results using an alternative measure of capital computed using the perpetual inventory method. For this measure, we use the gross value of property, plants, and equipment as our initial stock. We then measure net investment as the difference in the net capital stock—*PPENT* in Compustat—over consecutive periods and deflate this over time using the investment price deflator to compute the capital stock over time.
- 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 quintiles, we do so for each year and within each 3-digit NAICS industry. We also experimented with the classification of industries based on SIC codes used in Baqaee and Farhi (2020b) and obtained very similar results.
- Sample definition and trimming: following De Loecker, Eeckhout and Unger (2020), we trim the sample by removing firms in the bottom 5th and top 5th percentiles of the *COGS*-to-*SALES* distribution. In addition, following Baqaee and Farhi (2020b), we exclude firms in farm and agriculture, construction, real estate, finance, and utilities from our markup and labor share calculations in Figures 3 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.

S.1.2 Estimation approach and details

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^k_{tf} and the output-to-variable-input elasticity ε^v_{tf} 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:[35]

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

$$\ln y^R_{tf} = z^R_{tf} + \varepsilon^{Rv}_{tc(f)} \cdot \ln v_{tf} + \varepsilon^{Rk}_{tc(f)} \cdot \ln k_{tf} + \epsilon_{tf},$$

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

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

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

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 $c(f)$.

[35] An alternative approach to estimating markups assumes constant returns to scale (as we do) and 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.

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

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

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

$$\text{(A1)} \qquad \varepsilon^{v}_{tf} = \varepsilon^{Rv}_{tc(f)} \Big/ \left(\varepsilon^{Rv}_{tc(f)} + \varepsilon^{Rk}_{tc(f)}\right) \qquad \varepsilon^{k}_{tf} = \varepsilon^{Rk}_{tc(f)} \Big/ \left(\varepsilon^{Rk}_{tc(f)} + \varepsilon^{Rk}_{tc(f)}\right).$$

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 $k_{t+1,f} = \pi(k_{tf}, z^{R}_{tf})$, 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 do not observe prices, such that the usual estimation procedure yields revenue elasticities, not the quantity elasticities that are relevant for computing markups (see Bond et al., 2021). Under Assumption A5 we can recover output elasticities from revenue elasticities using (A1). 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.

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

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

Here $\theta^{h}_{tc(f)}$ is a parametrization for a flexible function h that might vary over time and between groups of firms. For any pair of revenue elasticities $\varepsilon^{Rv}_{tc(f)}$ and $\varepsilon^{Rk}_{tc(f)}$, 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

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

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

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

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. Thus, our estimation provides output elasticities that vary over time, by industry, and by quintiles of firm size in each industry. This represents a significant deviation from previous papers which assume that all firms in a given industry share the same output elasticities.

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.

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 S1 plots the behavior of the resulting output elasticities over time by firm size quintile. Figure S2 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 S3 plots the behavior of the resulting output elasticities over time by firm size quintile. Figure S4 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

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

Define $\upsilon_{tf} = z^R_{tf} + \epsilon_{tf}$. 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

$$\upsilon_{tf} = \ln y^R_{tf} - \varepsilon^{Rv}_{tc(f)} \cdot \ln v_{tf} - \varepsilon^{Rk}_{tc(f)} \cdot \ln k_{tf}.$$

Then we estimate the linear model

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

and base estimation on the moment conditions

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

Figure S5 plots the behavior of the resulting output elasticities over time by firm size quintile. Figure S6 reports the contribution of within-firm changes in markups and between-firm reallocation to (percent) changes in the labor share.

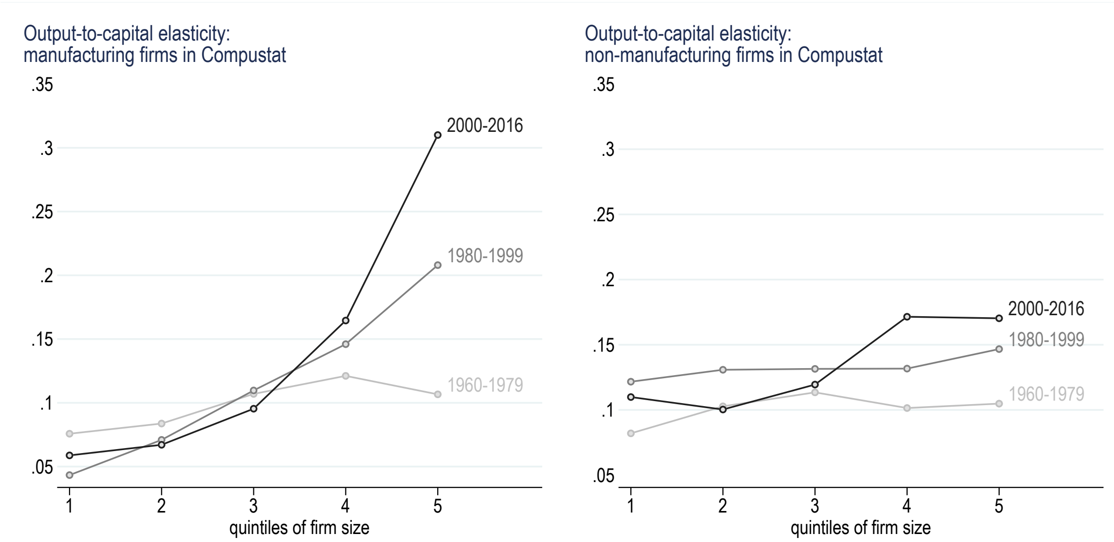

FIGURE S1: OUTPUT-TO-CAPITAL ELASTICITIES FOR COMPUSTAT FIRMS ESTIMATED USING A CUBIC PARAMETRIZATION OF g AND h. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms. Firm-level elasticities are estimated using a cubic parametrization for g and h, as explained in Appendix S.1.

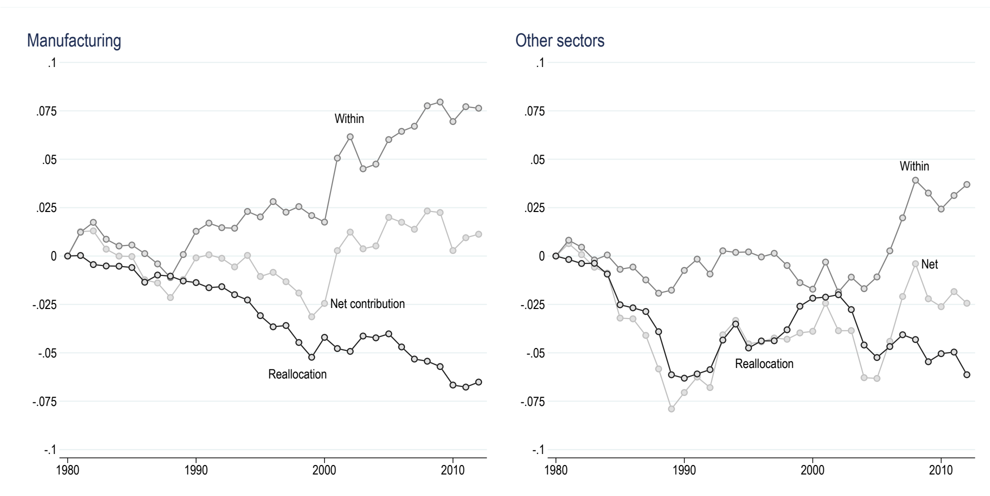

FIGURE S2: DECOMPOSITION OF THE CONTRIBUTION OF WITHIN-FIRM CHANGES IN MARKUPS AND BETWEEN-FIRM REALLOCATION TO (PERCENT) CHANGES IN THE LABOR SHARE. See the main text for details on this decomposition. Firm-level markups are estimated using a cubic parametrization for g and h, as explained in Appendix S.1. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other economic sectors.

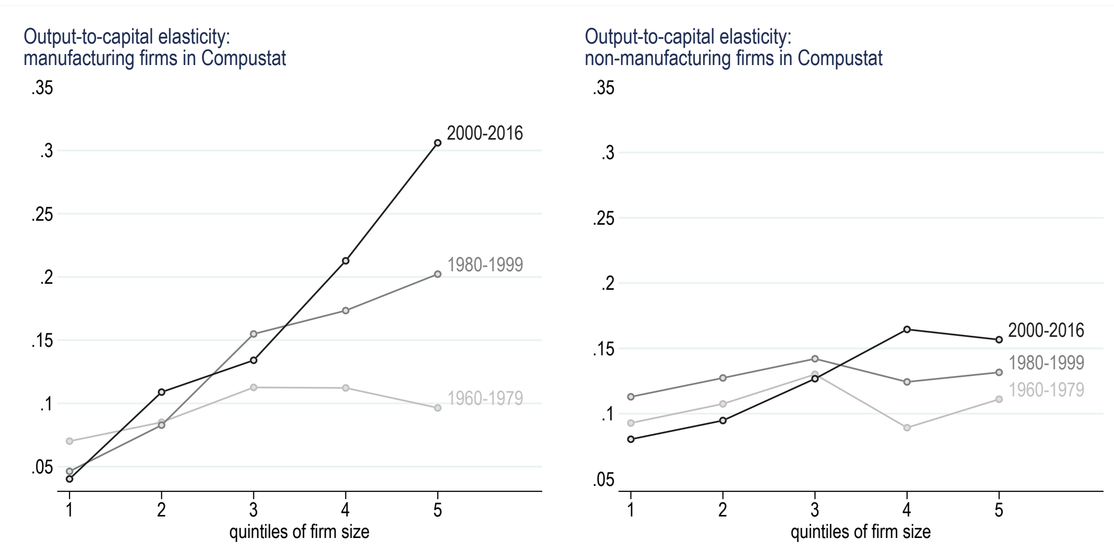

Figure S3: Output-to-capital elasticities for Compustat firms estimated under the assumption that there are no ex-post shocks. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms. Firm-level elasticities are estimated under the assumption of no ex-post shocks, as explained in Appendix S.1.

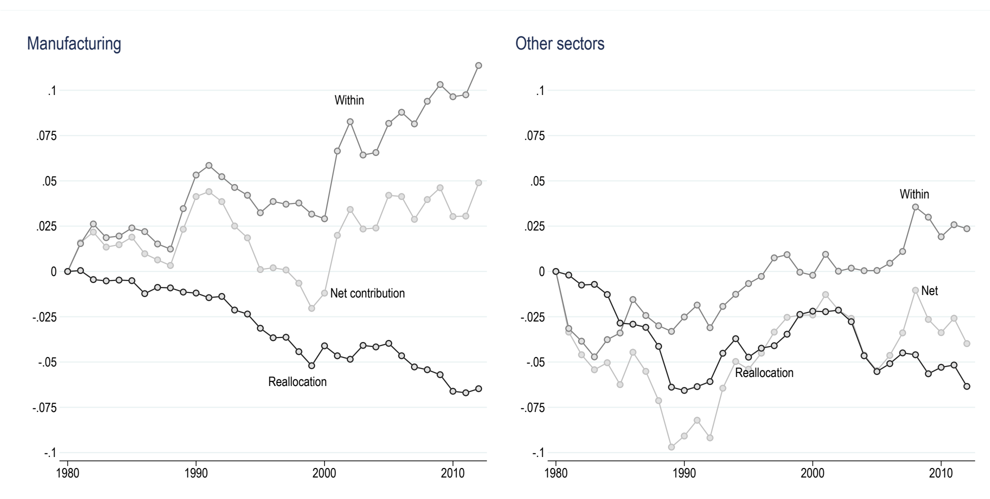

Figure S4: Decomposition of the contribution of within-firm changes in markups and between-firm reallocation to (percent) changes in the labor share. See the main text for details on this decomposition. Firm-level markups are estimated under the assumption of no ex-post shocks, as explained in Appendix S.1. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other economic sectors.

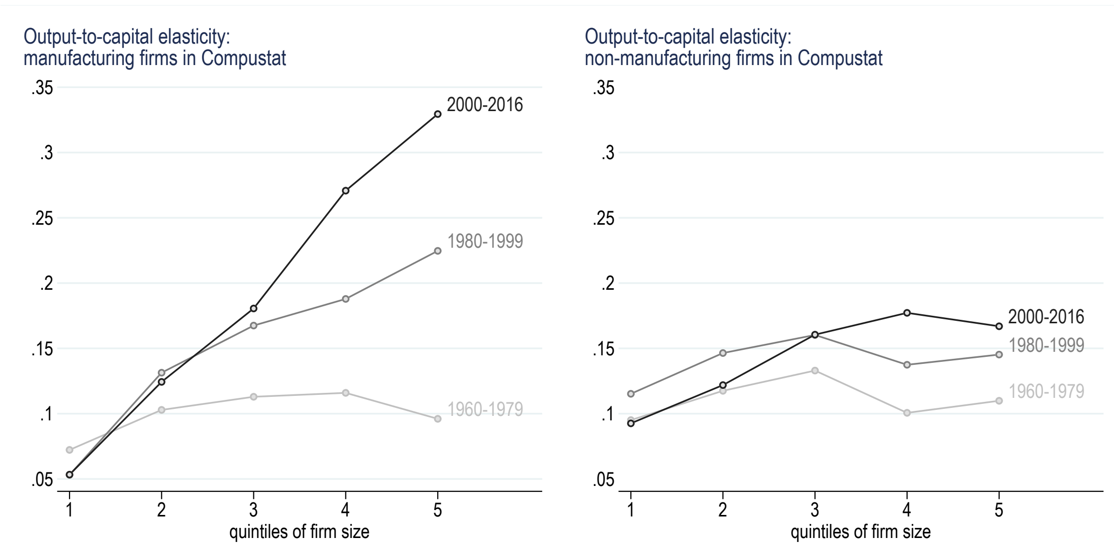

Figure S5: Output-to-capital elasticities for Compustat firms estimated under the assumption that productivity follows a linear Markov process. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms. Firm-level elasticities are estimated under the assumption that productivity follows a linear Markov process, as explained in Appendix S.1.

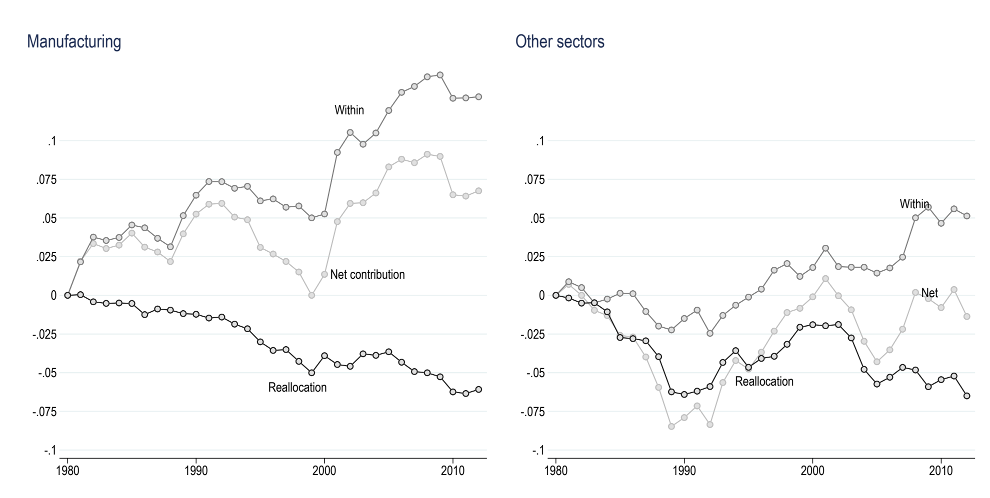

Figure S6: Decomposition of the contribution of within-firm changes in markups and between-firm reallocation to (percent) changes in the labor share. See the main text for details on this decomposition. Firm-level markups are estimated under the assumption that productivity follows a linear Markov process, as explained in Appendix S.1. The left panel provides the decomposition for manufacturing firms in Compustat. The right panel provides the decomposition for Compustat firms in other economic sectors.

S.1.3 Implications for the average markup in the economy as a whole

Figure S7 plots the implied time series for the economy-wide aggregate markup, computed as a sales-weighted *harmonic* mean of firm-level markups. Our estimates for markups suggest that they have been stable over time at around 1.2. This is in line with our quantitative exercise, which points to a modest increase in markups.

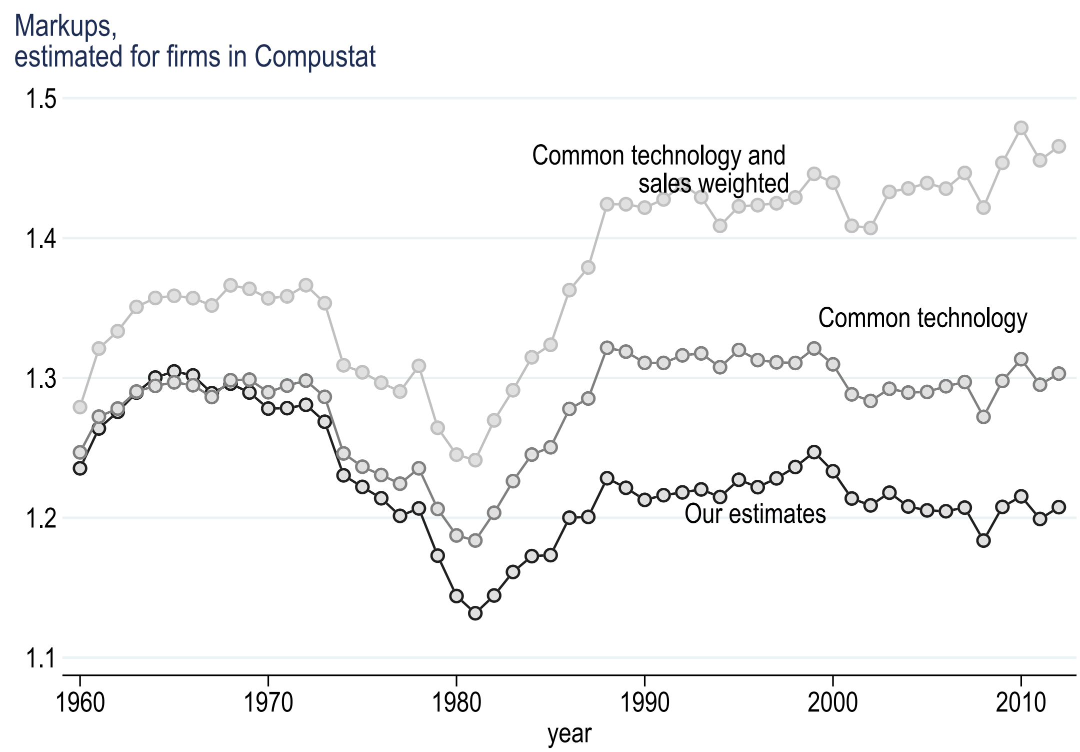

FIGURE S7: MARKUPS. The figure presents the aggregate markup for firms in Compustat. Our estimates are obtained as as a sales-weighted *harmonic* mean of firm-level markups. The figure also reports the aggregate markup that would result under the assumption of common output elasticities across firms in the same industry, and a version of these estimates that aggregates firms' markups using a sales-weighted arithmetic mean.

For comparison, we provide an alternative estimate of the aggregate markup obtained under the assumption that all firms in an industry operated technologies with the same capital intensity (as opposed to letting it vary by 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 aligns with the harmonic-mean estimates in Edmond, Midrigan and Xu (2022). We also 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, which coincide with the series in De Loecker, Eeckhout and Unger (2020). Despite its increasing trend over time, this series is inappropriate for understanding the contribution of markups to the decline in the labor share because it ignores differences in technology across firm-size classes and uses the wrong weights for aggregation.

S.1.4 Additional evidence from Compustat

This section provides additional descriptive statistics from Compustat that support the notion that large firms operate more capital-intensive technologies. In what follows, we estimate regression models of the form

$$\ln y_{tfi} = \alpha_{ti} + \beta_{tc(f)} + \varepsilon_{tfi}, \tag{A2}$$

where we explain different measures for the capital intensity y_{tfi} of firm f in industry i at time t as a function of industry and year fixed effects (the α_{ti}) and size class dummies ($\beta_{tc(f)}$) that are allowed to vary over time. In particular, we estimate different size-class dummies for the periods of 1960–1980, 1980–2000 and 2000–2016. We treat firms in the smallest size class of an industry as the excluded category and report estimates weighted by firm sales.

Figure S8 plots estimates of equation (A2) for firms' investment rates, defined as their investment ($CAPX$ in Compustat) normalized by variable cost (top panel), employment (middle panel), and sales (bottom panel). The left panel provides estimates for manufacturing firms and the right panel for firms outside of manufacturing. For the 1980–2000 and 2000–2016 period, the largest manufacturing firms in each industry have had investment rates 60–140 log points higher than those of the smallest firms. Outside of manufacturing, the difference is less pronounced, with the largest firms having 10–90 log points higher investment rates than the smallest firms in their industries. In both cases, the gradient by size has become steeper over time.

Figure S9 plots estimates for firms' capital intensity, defined as their net capital stock ($PPENT$ in Compustat) normalized by variable cost (top panel), employment (middle panel), and sales (bottom panel). For the 1980–2000 and 2000–2016 period, the largest manufacturing firms in each industry had a 55 log point higher capital to variable cost ratio, a 120 log point higher capital per worker, and a 45 log point higher capital to sales ratio than the smaller firms in their industries. Here too, we see some evidence of the gradient by size becoming steeper over time, though the gradient and its rotation are less pronounced outside of manufacturing.

Finally, Figure S10 plots estimates for firms' reliance on capital services. Along a balanced growth path, the flow value of capital services used by a firm can be computed as[36]

$$\text{flow value of capital services} = (r - g) \cdot \text{net capital stock} + \text{capital expenditures.}$$

The figure provides estimates normalizing the flow value of capital services by variable costs (so that we get a measure of capital services relative to variable input services), employment, and sales (a measure of capital services in sales). In this exercise, we fix $r - g = 2.5\%$, which aligns with the calibration in Farhi and François (2018). For the 1980–2000 and 2000–2016 period, the largest manufacturing firms in each industry had a 70 log point higher reliance on capital services vs. variable input services when compared to the smallest firms in their industries. Here too, we see some evidence of the gradient by size becoming steeper over time, with the gradient and its rotation being less pronounced outside of manufacturing.

[36] In particular, suppose the firm faces no adjustment costs. Then the PDV of capital services equal the PDV of capital costs. The PDV of capital costs are $(1+r) \cdot$ net capital stock $+ \frac{1+r}{r-g} \cdot$ capital expenditure, which gives the cost of the initially installed capital and of financing it plus the PDV of capital expenditures. The flow value of capital services is $\frac{r-g}{1+r} \cdot$ PDV of capital costs and we get the formula in the text.

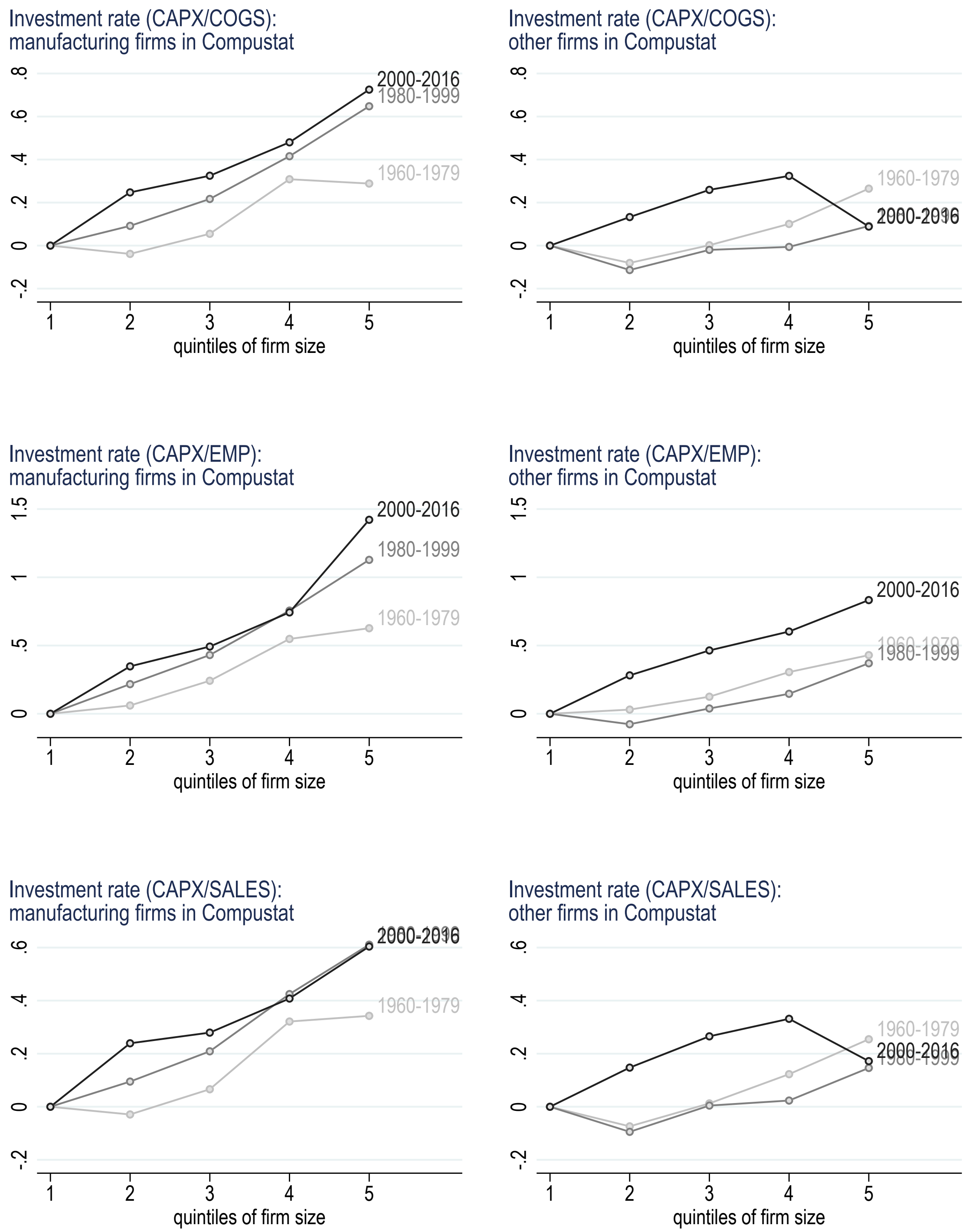

FIGURE S8: INVESTMENT RATES, COMPUSTAT. The figure presents estimates of the relative difference in investment rates by firm-size class using Compustat. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms.

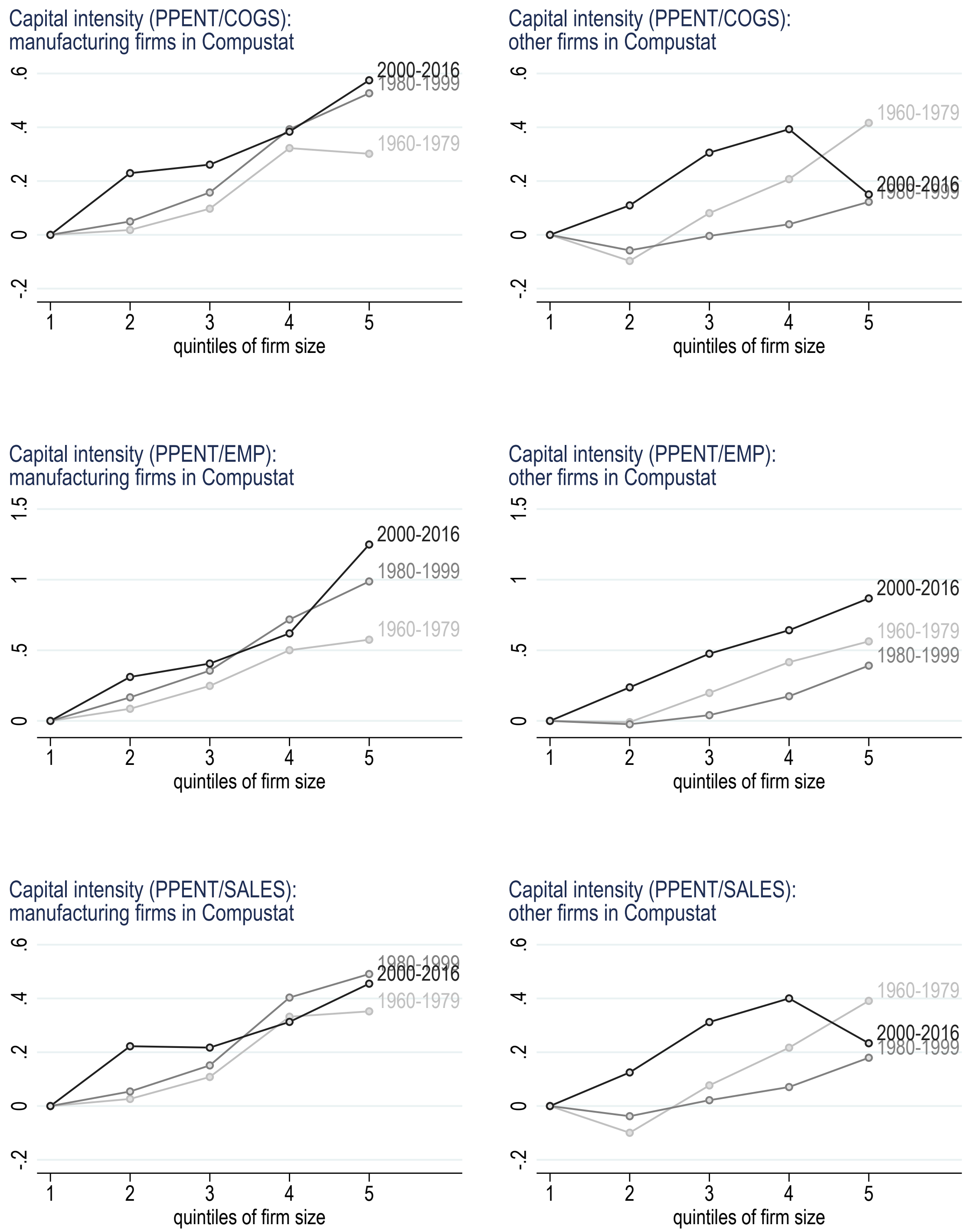

FIGURE S9: CAPITAL INTENSITY, COMPUSTAT. The figure presents estimates of the relative difference in capital intensity by firm-size class using Compustat. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms.

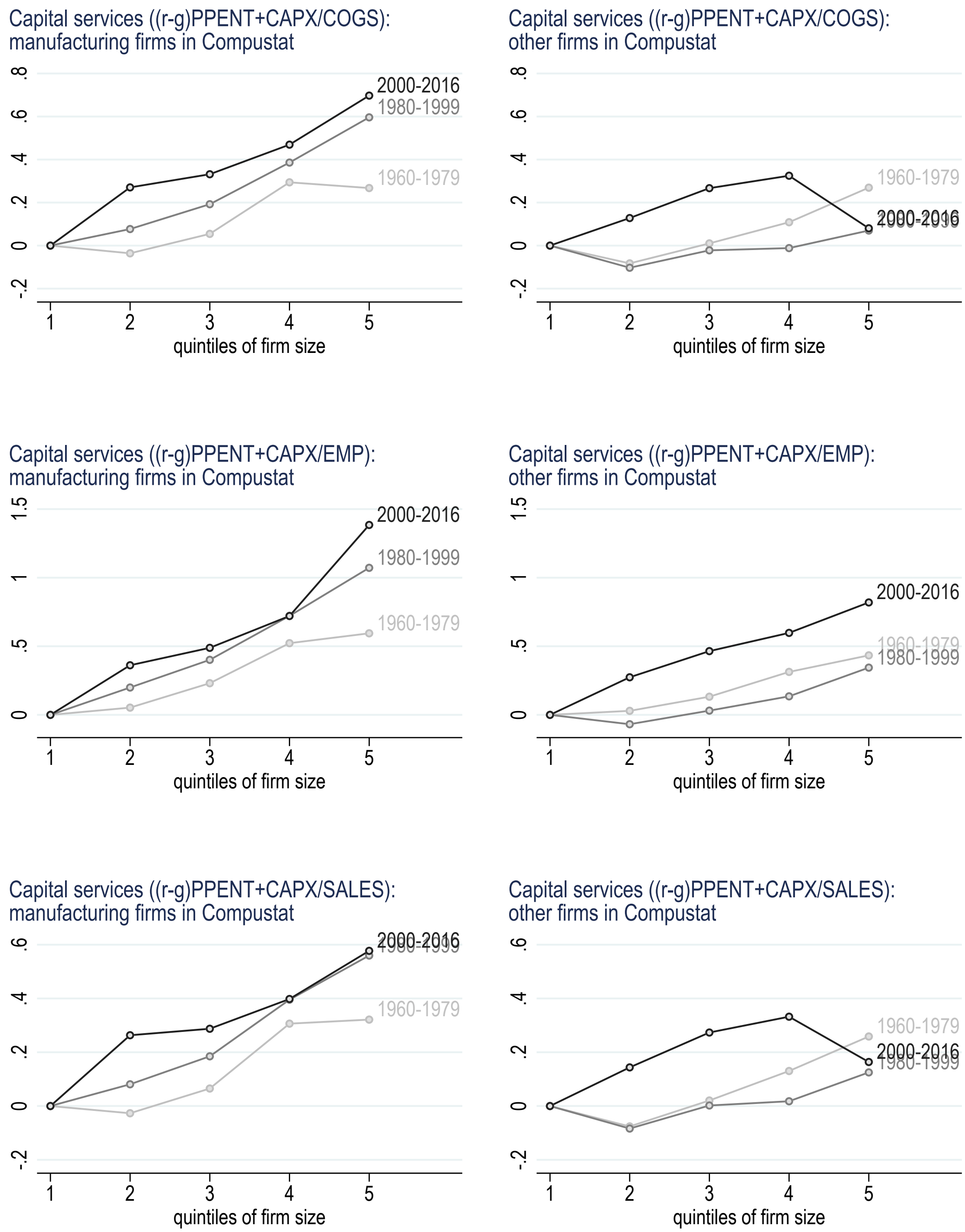

Figure S10: Capital services, Compustat. The figure presents estimates of the relative difference in capital services by firm-size class using Compustat. The left panel presents estimates for Compustat manufacturing firms. The right panel presents estimates for Compustat non-manufacturing firms.

S.2 Additional Data and Measurement Details

Historical behavior of labor shares: This section provides additional motivation for our focus on the 1982–2012 period. As a starting point, Figure S11 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*. As discussed in the main text, Figure S11 shows that payroll shares were constant or increasing up to 1982, and then started a sharp decline both in manufacturing, retail and wholesale. Payroll shares differ from labor shares in that they exclude compensation and self employment. But looking at payroll shares has the advantage of allowing us to go back further in time.

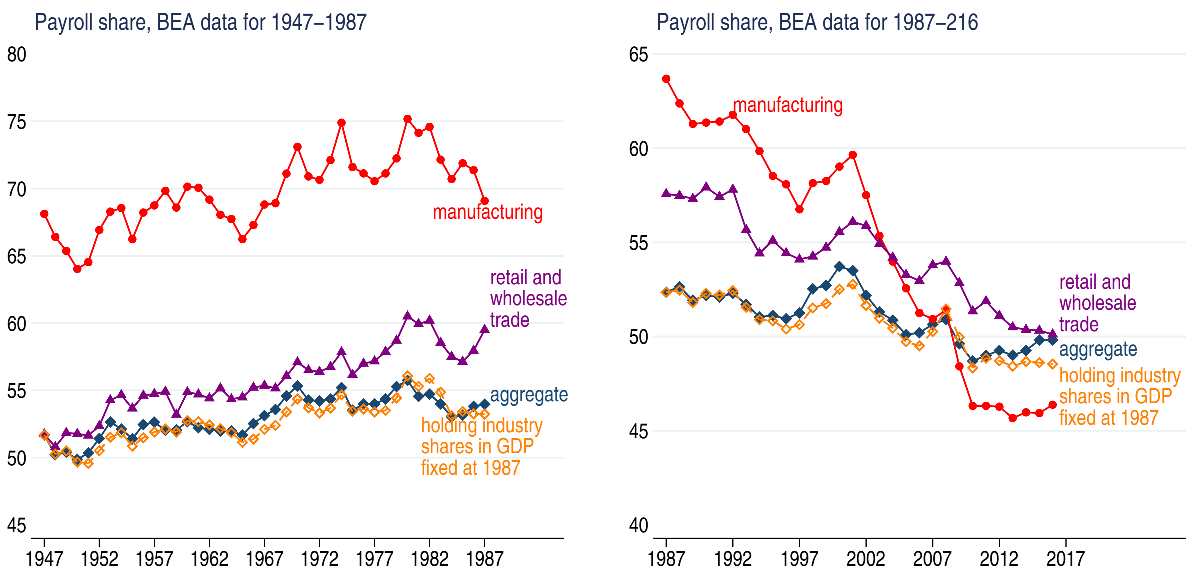

Figure S11: Payroll share in the US for 1947–2016. The figure plots the payroll share of value added, both for some specific sectors and the economy as a whole. Data from the BEA industry accounts. Industry definitions based on SIC in left panel, NAICS in right panel.

Labor shares (which also include non-wage compensation) are available starting in 1963 from the BEA-BLS integrated industry-level production account. Figure S12 confirms that labor shares 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.

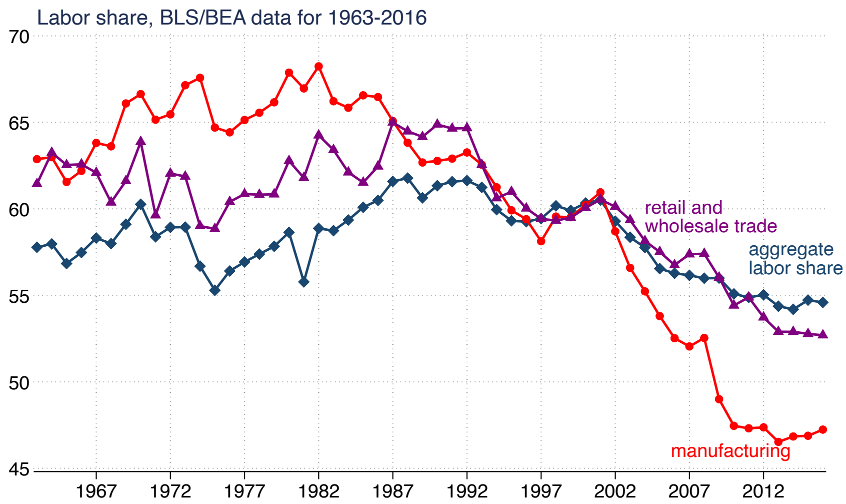

FIGURE S12: LABOR SHARE IN THE US FOR 1963–2016 The figure plots the labor share of value added, both for some specific sectors and the economy as a whole. Data from the BEA-BLS integrated industry-level production account.

Measures of capital prices: We create measures of quality-adjusted capital prices at the sectoral level building on DiCecio (2009), Cummins and Violante (2002), and Gordon (1990).

In the first step, we obtain data on nonresidential asset prices and quantities from the BEA Fixed Asset Tables. These data cover 39 types of equipment, 32 types of structures, 3 types of software, and 22 types of other intellectual property products. We exclude other intellectual property products from our analysis since these would be considered part of the fixed cost of adopting new technologies in our model. The data cover the period from 1947 to 2020 and include information on investment at current nominal prices, investment at constant 2012 prices, stocks, and depreciation.

Using these series, we construct a price index for each detailed asset a as

$$p_{a,t} = \frac{\text{investment at current nominal prices}_{a,t}}{\text{investment at constant 2012 prices}_{a,t}}.$$

We let $\Delta \ln p_{a,t}$ denote the percent change in asset prices between time t and $t+1$.

Our second step involves adjusting the BEA prices for quality. We follow the work by DiCecio (2009) and Cummins and Violante (2002). These authors use the series for quality-adjusted investment prices from Gordon (1990), and which covered the postwar period up to 1983, and extend it from 1947 to 2011. Cummins and Violante (2002) estimate a statistical model explaining Gordon's quality-adjusted price indices as a function of those by the BEA/NIPA, their lags, and time trends. They then extrapolate this model to produce quality-adjusted price indices for 1947–2000. DiCecio (2009) follows the same procedure and creates an updated series up to 2011 for equipment and software. We use the estimates from DiCecio (2009) on the quality adjusted price of equipment and software, obtained via FRED (variable code PERICD), and denote the percent change in the quality-adjusted price of software and equipment as $\Delta \ln p^*_{t,\text{E\&S}}$. We then compute a user-cost weighted price index for software and equipment from the BEA data using a Törnqvist index

$$\Delta \ln p_{t,\text{E\&S}} = \sum_{a \in \text{E\&S}} \frac{1}{2} \cdot \left(s^{\text{E\&S}}_{t+1,a} + s^{\text{E\&S}}_{a,t} \right) \cdot \Delta \ln p_{a,t},$$

where $s^{\text{E\&S}}_{a,t}$ denotes the share of asset a in equipment and software capital services.[37] The implied quality adjustment for equipment and software is therefore equal to

$$\Delta \text{quality adjustment}_{\text{E\&S},t} = \Delta \ln p_{E\&S,t} - \Delta \ln p^*_{\text{E\&S},t}.$$

In the BEA data, the price of equipment and software declined by an average of 3.1% per year in 1980-2011. The quality-adjusted series from DiCecio (2009) shows a decline on 5.7% per year, which implies an improvement in the quality of equipment and software of 2.6% per year.

We then compute a quality adjusted series for the detailed equipment and software products in the BEA data as

$$\Delta \ln p^*_{a,t} = \Delta \ln p_{a,t} - \Delta \text{quality adjustment}_{\text{E\&S},t}.$$

This assumes a common quality adjustment for all types of software and equipment. For structures, we do not perform quality adjustment.

In the third step, we account for changes in taxation using the estimates in Acemoglu, Manera and Restrepo (2020) of effective taxes on equipment, software, and structures. These authors estimate that the effective tax on equipment decreased from 12.4% to 4.7% during 1981–2018, the effective tax on software decreased from 14.6% to 4.7% during 1981–2018, and the effective tax on structures increased from 8.3% to 9% during 1981–2018. These changes in taxes imply a further reduction in the cost of producing tasks with capital of

[37] We compute capital services derived from an asset a as

$$\text{capital services}_{a,t} = (r + \delta_a - \Delta \ln p_{a,t}) \cdot \text{stock asset}_{a,t},$$

where we take a required rate of return $r = 4\%$, and use the depreciation rate and change in capital prices from the BEA.

close to 10% during the whole 1981–2018 period.

In the fourth step, we compute a measure for the relative price of capital by asset, $\Delta \ln p^{*,r}_{a,t}$ by taking our quality-adjusted price indices adjusted for taxes and subtracting changes in the BEA price of consumption expenditures index.

In the final step, we construct a sector-specific measure of capital prices using a user-cost weighted Törnqvist index

$$\Delta \ln p_{i,t} = \sum_{a} \frac{1}{2} \cdot \left(s^{i,k}_{a,t+1} + s^{i,k}_{a,t} \right) \cdot \Delta \ln p^{*,r}_{a,t},$$

where $s^{i,k}_{a,t}$ denotes the share of asset a in total capital services in sector i, computed also from the industry-level version of the BEA Fixed Asset Tables. This index provides the average decline in capital prices used in sector i over the 1980–2011 period.

Our resulting sectoral price indices imply that the average price of capital used in manufacturing declined by 108 log points from 1980 to 2011. For retail, the average decline was of 86 log points, for utilities 68 log points, and for wholesale of 159 log points. These differences across sectors reflect the different bundles of capital goods used, with manufacturing and wholesale investing more heavily in equipment and software.

Supplementary Material References

Acemoglu, D., A. Manera, and P. Restrepo (2020): "Does the US Tax Code Favor Automation?," *Brookings Papers on Economic Activity*, 2020(1), 231–300.

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83(6), 2411–2451.

Baqaee, D. R., and E. Farhi (2020): "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 135(1), 105–163.

Bond, S., A. Hashemi, G. Kaplan, and P. Zoch (2021): "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data," *Journal of Monetary Economics*, 121, 1–14.

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.

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.

DiCecio, R. (2009): "Sticky wages and sectoral labor comovement," *Journal of Economic Dynamics and Control*, 33(3), 538–553.

Edmond, C., V. Midrigan, and D. Y. Xu (2022): "How Costly Are Markups?," *Journal of Political Economy*, forthcoming.

Farhi, E., and G. François (2018): "Accounting for Macro-Finance Trends: Market Power, Intangibles, and Risk Premia," *Brookings Papers on Economic Activity*, p. 147.

Gordon, R. J. (1990): *The Measurement of Durable Goods Prices*. University of Chicago Press.

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

Olley, G. S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1297.