ONLINE APPENDICES (NOT INTENDED FOR PUBLICATION)

Appendix A MODEL OF SUPERSTAR FIRMS[56]

In this Appendix we derive conditions under which changes in the market environment will affect the equilibrium labor share. We derive three key results. First, larger firms will have lower labor shares. Second, an increase in the "toughness" of the market (e.g. because of increased market size due to globalization or greater competition) will reallocate output towards low labor share firms, which will in turn tend to lower the aggregate labor share (a "between-firm effect"). Third, the increase in market toughness will *increase* individual firms' labor shares as markups fall, a within-firm effect. The net effect of an increase in market toughness on aggregate industry-wide labor share will depend on the balance of these two forces, and that is in turn determined by the underlying productivity distribution. When the underlying probability density function of firm productivity draws is log-linear (e.g. Pareto) the two forces will perfectly counterbalance and the labor share will be unchanged. When this pdf is log convex, the aggregate labor share will fall when markets gets tougher. The opposite is true for log-concavity.

Appendix A.1 Basic Environment

Consider an industry with monopolistic competition and firm-level heterogeneity in productivity (z). Ω denotes the set of differentiated varieties. Labor, V, is the only factor of production, cost functions are linear (implying a constant marginal cost, c), M denotes market size and w denotes the wage.

Demand Structure

Individual demand for any good $\omega \in \Omega$ takes the form

$$q(p_\omega) = p_\omega^{-\sigma} d\left(A p_\omega\right) \tag{9}$$

where p_ω is the price of good ω; σ is an exogenous preference parameter; and A is an endogenous demand shifter. Each firm produces one good/variety. In addition, we assume that $d(\cdot)$ is such that:

- there exists a "choke price" $\bar{p}$ such that $d(p) = 0$ for all $p \geq \bar{p}$
- $d(p) > 0$, $d'(p) < 0$, $d \ln d(p)/d \ln p < (\sigma - 1)$, and $d^2 \ln d(p)/d\left(\ln p\right)^2 < 0$ ("Marshall's Second Law") for all $p < \bar{p}$

Examples of utility and expenditure functions satisfying equation (9) include the Additively Separable Utility function, the Translog Expenditure Function, and the Quadratic Utility Function used by *inter alia* Melitz and Ottaviano (2008). A key feature of these demand systems is that they obey Marshall's Second Law of Demand according to which the absolute elasticity of demand is lower for higher levels of consumption (lower levels of price). For example, consider the Quadratic

[56] We are extremely grateful to Arnaud Costinot for extensive help with this Appendix, which is largely based on earlier versions of Arkolakis et al (2018).

Utility Function:

$$U = q^0 + \alpha \int_{\omega\in\Omega} q_\omega d\omega - \frac{1}{2}\gamma \int_{\omega\in\Omega} (q_\omega)^2 d\omega - \frac{1}{2}\eta \left(\int_{\omega\in\Omega} q_\omega d\omega \right)^2$$

where $\alpha, \gamma, \eta > 0$ and q^0 and q_ω represent the individual consumption levels of the numeraire good and variety ω. γ indexes the degree of product differentiation between varieties (when $\gamma = 0$ the varieties are perfect substitutes). The inverse demand (when $q_\omega > 0$) for each variety is linear:

$$q(p_\omega) = p_\omega \left[\frac{1}{p_\omega A} - \frac{1}{\gamma} \right],$$

where

$$A = \left[\frac{\alpha}{\eta N + \gamma} + \frac{\eta}{\gamma} \frac{\int_{\omega\in\Omega|q_\omega>0} p_\omega d\omega}{\eta N + \gamma} \right]^{-1}.$$

and $N \equiv \int_{\omega\in\Omega} d\omega$ is the number of consumed varieties (where $q_\omega > 0$). This implies that $\sigma = -1$ and $d(p_\omega A) = \left(\frac{1}{p_\omega A} - \frac{1}{\gamma} \right)$.

Note that although most classical demand functions are consistent with Marshall's Second Law, there are exceptions.[57]

Entry, Pricing and markups

Firms choosing to enter must bear an entry cost $\kappa > 0$. After fixed entry costs have been paid, firms receive a random productivity draw z from a commonly known distribution with pdf $\lambda(z)$. For a firm with productivity z, the cost of producing one unit of a good is given by $c = 1/z$. Consider the firm producing good ω. Let c_ω denote its constant marginal cost. The firm chooses its price p_ω in order to maximize profits

$$(p_\omega - c_\omega)\, q(p_\omega)$$

taking as given the demand shifter A. The associated first-order condition is

$$q(p_\omega) + (p_\omega - c_\omega)\, q'(p_\omega) = 0 \tag{10}$$

so that

$$\frac{p_\omega - c_\omega}{p_\omega} = -\frac{1}{p_\omega} \frac{q(p_\omega)}{q'(p_\omega)} = -\frac{1}{\varepsilon(p_\omega)}, \tag{11}$$

where $\varepsilon(p_\omega) \equiv d \ln q(p_\omega) / d \ln p_\omega$ is the demand elasticity. Using equation (9) we can express this elasticity as

$$q(p_\omega) = p_\omega^{-\sigma} d(Ap_\omega)$$

$$\varepsilon(p_\omega) = Ap_\omega d'(Ap_\omega) / d(Ap_\omega) - \sigma$$

Letting $m_\omega \equiv p_\omega / c_\omega$ denote the markup, we obtain

$$m_\omega = m(Ap_\omega), \tag{12}$$

[57] See Mrazova and Neary (2017) for a general discussion. For example, under Dixit-Stigliz CES preferences, the demand elasticity is constant so $d^2 \ln d(p) / d(\ln p)^2 = 0$.

where

$$m(p) \equiv \frac{\sigma - pd'(p)/d(p)}{\sigma - 1 - pd'(p)/d(p)}. \tag{13}$$

Since labor is the only factor, unit cost is $c_\omega = wV/q$. So from the markup definition, the share of labor in revenues is simply the inverse of the markup:

$$S_\omega \equiv \frac{wV}{p_\omega q_\omega} = \frac{c_\omega}{p_\omega} = \frac{1}{m_\omega}. \tag{14}$$

In order to see how the labor share changes, we need to characterize the determination and distribution of markups.

Appendix A.2 Firm level results

Claim 1: *Prices are strictly increasing with marginal costs.*

Proof: Note that from differentiating the markup definition and rearranging, we also have

$$\frac{\partial p(c_\omega, A)}{\partial c_\omega} = \frac{m(p_\omega)}{1 - m'(p_\omega)c_\omega} > 0$$

which is positive since $m(p_\omega) > 0$ and $m'(p_\omega) < 0$

Claim 2: *Markups are strictly decreasing with marginal costs.*

Proof: By equation (10), we know that

$$m(p) = 1 - \frac{1}{Apd'(Ap)/d(Ap) - \sigma + 1} \tag{15}$$

Since $d^2 \ln d(p)/d(\ln p)^2 < 0$, we therefore have $m'(p) < 0$. Since prices are increasing with marginal costs, by Claim 1, markups are decreasing with marginal costs.

Claim 3: *There exists a cutoff $c^* = \bar{p}/A$ such that firms produce if and only $c_\omega \leq c^*$. Furthermore, the markup for a firm with marginal cost c^* is equal to one.*

Proof: Since prices are strictly increasing with marginal costs and demand is zero if $p > \bar{p}/A$, there exists a cutoff c^* such that firms produce if and only if $c_\omega \leq c^*$. At the cutoff c^*, the firm faces zero demand and charges $\bar{p}/A$. Thus, given equation (15), the firm has a markup equal to one, hence

$$c^* = \bar{p}/A. \tag{16}$$

Recalling that $M =$ market size, we can now derive a number of key objects of interest in terms of relative costs.

Claim 4: *Prices (p), markups (m), total output (Q), total sales (r), and total profits (π) can be expressed as*

$$p(\ln c_\omega, \ln c^*) = e^{\ln c_\omega} f(\ln c_\omega - \ln c^*) \tag{17}$$
$$m(\ln c_\omega, \ln c^*) = f(\ln c_\omega - \ln c^*) \tag{18}$$
$$Q(\ln c_\omega, \ln c^*) = Me^{-\sigma \ln c_\omega} h(\ln c_\omega - \ln c^*) \tag{19}$$
$$r(\ln c_\omega, \ln c^*) = Me^{(1-\sigma)\ln c_\omega} f(\ln c_\omega - \ln c^*) h(\ln c_\omega - \ln c^*) \tag{20}$$
$$\pi(\ln c_\omega, \ln c^*) = Me^{(1-\sigma)\ln c_\omega} \left[f(\ln c_\omega - \ln c^*) - 1\right] h(\ln c_\omega - \ln c^*) \tag{21}$$

where $f(x)$ is implicitly defined as the solution in y of the equation

$$y = m(\bar{p}ye^{x})$$

and where $h(x)$ is defined by

$$h(x) \equiv (f(x))^{-\sigma}\, d(e^{x}f(x)\bar{p}).$$

Proof: By definition of the markup and equation (12), we know that

$$p_\omega = c_\omega m(Ap_\omega)$$

which can be rearranged as

$$\frac{p_\omega}{c_\omega} = m\left(\bar{p}\frac{p_\omega}{c_\omega}\frac{c_\omega}{c^*}\right),$$

given equation (16). Equation (17) directly derives from this expression. The other equations can be established by simple substitutions. Thus we can state:

Proposition 1 *Large firms will have lower labor shares.*
Proof. *Low marginal cost firms will be larger (they have lower prices and higher demand). By claim 2 they will also have higher markups and labor share is the reciprocal of the markup (equation 14).*

Appendix A.3 Industry Level Results

One can think of c^*, which corresponds to the maximum feasible break-even price, as a measure of the toughness of the market. The lower is c^*, the tougher the market. How will changes in the toughness of the market affect the distribution of firm markups and the aggregate mark-up, and so the labor share?

Distribution of markups

Let $\Phi(m, c^*) = \Pr\{X \leq m | c \leq c^*\}$ denote the distribution of markups for a given level of toughness of the market. By Bayes' rule, this can be rearranged as

$$\begin{aligned}\Phi(m, c^*) &= \frac{\Pr\{f(\ln c - \ln c^*) \leq m, \ln c \leq \ln c^*\}}{\Pr\{\ln c \leq \ln c^*\}} \\ &= \frac{\Pr\left\{\ln c^* + f^{-1}(m) \leq \ln c \leq \ln c^*\right\}}{\Pr\{\ln c \leq \ln c^*\}},\end{aligned}$$

where the second inequality uses the fact that $f' < 0$.

Given our choice of numeraire, we know that $\ln c = -\ln z$. So letting $\ln c^* = -\ln z^*$, we can rearrange the previous expression as

$$\begin{aligned}\Phi(m, \ln z^*) &= \frac{\int_{\ln z^*}^{\ln z^* - f^{-1}(m)} \lambda(u)du}{1 - \Lambda(\ln z^*)} = \frac{\Lambda\left[\ln z^* - f^{-1}(m)\right] - \Lambda(\ln z^*)}{1 - \Lambda(\ln z^*)} \\ &= \frac{\Lambda\left[\ln z^* - f^{-1}(m)\right] - 1}{1 - \Lambda(\ln z^*)} + 1\end{aligned}$$

where λ and Λ are the pdf and cdf of log-productivity, respectively.

Thus the conditional density of markups (ϕ is the the PDF of Φ) is given by

$$\phi(m, \ln z^*) = \frac{-f^{-1\prime}(m)\,\lambda\left[\ln z^* - f^{-1}(m)\right]}{1 - \Lambda(\ln z^*)} \implies$$

$$\ln \phi(m, \ln z^*) = \ln\left[-f^{-1\prime}(m)\right] + \ln\left\{\lambda\left[\ln z^* - f^{-1}(m)\right]\right\} - \ln\left[1 - \Lambda(\ln z^*)\right] \ .$$

Notice that the above implies that

$$\frac{\partial^2 \ln \phi(m, \ln z^*)}{\partial m \partial \ln z^*} = \frac{\partial^2 \ln\left\{\lambda\left[\ln z^* - f^{-1}(m)\right]\right\}}{\partial m \partial \ln z^*}$$

Since $\frac{\partial \ln \lambda[\ln z^* - f^{-1}(m)]}{\partial \ln z^*}$ is a function of $\ln z^* - f^{-1}(m)$, it is immediate that we have

$$\frac{\partial \ln\left\{\lambda\left[\ln z^* - f^{-1}(m)\right]\right\}}{\partial m \partial \ln z^*} \quad = \left[\frac{\partial\left(-f^{-1}(m)\right)}{\partial m}\right]\left(\frac{\partial^2 \ln \lambda\left[\ln z^* - f^{-1}(m)\right]}{\partial \ln (z^*)^2}\right) \tag{22}$$

Since $f' < 0$, $-f^{-1}(\cdot)$ is increasing in m, the first term on the right hand side of equation (22), $\left[\frac{\partial\left(-f^{-1}(m)\right)}{\partial m}\right]$, is positive.

Thus the sign of $\frac{\partial^2 \ln \phi(m, \ln z^*)}{\partial m \partial \ln z^*}$ is the same as the sign of $\frac{\partial^2 \ln \lambda\left[\ln z^* - f^{-1}(m)\right]}{\partial \ln (z^*)^2}$.

Accordingly, we have:

- ϕ log-supermodular in $(m, \ln z^*)$ if λ log-convex;
- ϕ log-submodular in $(m, \ln z^*)$ if λ log-concave;
- ϕ multiplicatively separable in $(m, \ln z^*)$ if λ log-linear.

Log-supermodularity implies the monotone likelihood ratio property (MLRP, cf. Costinot 2009). Thus, we can state the following proposition.

Proposition 2 *Consider $c^{*\prime} \leq c^*$. Then:*

- $\Phi(\cdot, c^*) \prec_{mlrp} \Phi(\cdot, c^{*\prime})$ *if λ log-convex;*
- $\Phi(\cdot, c^*) \succ_{mlrp} \Phi(\cdot, c^{*\prime})$ *if λ log-concave;*
- $\Phi(\cdot, c^*) = \Phi(\cdot, c^{*\prime})$ *if λ log-linear.*

Since dominance in terms of MLRP is stronger than dominance in terms of First Order Stochastic Dominance, we obtain the following corollary.

Corollary 1. *In tougher markets, the average labor share is lower (and markup is higher) if λ is log-convex, higher if λ is log-concave, and the same if λ is log-linear.*

Share of aggregate profits How do the previous results regarding the distribution of markups translate into predictions about the share of aggregate profits? Given equations (20) and (21), we can express aggregate revenues and aggregate profits as

$$\begin{aligned} R &= NM \int_{\ln z^*}^{\infty} e^{(\sigma-1)u} f(\ln z^* - u) h(\ln z^* - u) \lambda(u) du \\ \Pi &= NM \int_{\ln z^*}^{\infty} e^{(\sigma-1)u} \left[f(\ln z^* - u) - 1 \right] h(\ln z^* - u) \lambda(u) du \end{aligned}$$

where N is the number of firms and M is market size. Note that we have multiplied both integrals by M to go from individual to aggregate demand.

Changing variable, $v \equiv \ln z^* - u$, we obtain

$$\begin{aligned} R &= NM \int_{-\infty}^{0} (z^*)^{\sigma-1} e^{(1-\sigma)v} f(v) h(v) \lambda(\ln z^* - v) dv \\ \Pi &= NM \int_{-\infty}^{0} (z^*)^{\sigma-1} e^{(1-\sigma)v} \left[f(v) - 1 \right] h(v) \lambda(\ln z^* - v) dv \end{aligned}$$

Let us introduce $b(v, \delta) \equiv NMe^{(1-\sigma)v} \left[f(v) + \delta \right] h(v)$. By construction, we have

$$\frac{b(v,0)}{b(v,-1)} = \left[\frac{f(v)}{f(v)-1} \right]$$

which is increasing with v, since $f'(\nu) < 0$. Thus $b(v, \delta)$ is log-supermodular in (v, δ).

Now let us write

$$B(\ln z^*, \delta) = e^{(\sigma-1)\ln z^*} \int_{-\infty}^{0} b(v, \delta) \lambda(\ln z^* - v) dv$$

If λ is log-concave, then $\lambda(\ln z^* - v)$ is log-supermodular in $(\ln z^*, v)$. Since log-supermodularity is preserved by multiplication and integration, we have $B(\ln z^*, \delta)$ log-supermodular. This implies that if $\ln z^* \geq \ln z^{*\prime}$, then

$$\frac{B(\ln z^*, -1)}{B(\ln z^*, 0)} \leq \frac{B(\ln z^{*\prime}, -1)}{B(\ln z^{*\prime}, 0)}.$$

By construction we have

$$\begin{aligned} \Pi &= B(\ln z^*, -1) \\ R &= B(\ln z^*, 0) \end{aligned}$$

Thus the previous inequality implies that if λ is log-concave, then the share of aggregate profits is lower in tougher markets:

$$\left(\frac{\Pi}{R} \right)_{\ln z^*} \geq \left(\frac{\Pi}{R} \right)_{\ln z^{*\prime}}.$$

What if λ is log-convex? In this case, let us write

$$B(\ln z^*, \delta) = e^{(\sigma-1)\ln z^*} \int_{0}^{\infty} \widetilde{b}(u, \delta) \lambda(\ln z^* + u) du$$

where $\widetilde{b}(u, \delta) \equiv NMe^{-(1-\sigma)u} \left[f(-u) - \delta \right] h(u)$. Since λ is log-convex, $\lambda(\ln z^* + u)$ is log-supermodular in $(\ln z^*, u)$. Since $\widetilde{b}(u, \delta) \equiv b(-u, -\delta)$, $\widetilde{b}(u, \delta)$ is log-supermodular as well. Thus $B(\ln z^*, \delta)$ remains log-supermodular. But by construction, we now have:

$$\begin{aligned} \Pi &= B(\ln z^*, 1) \\ R &= B(\ln z^*, 0) \end{aligned}$$

Thus the log-supermodularity of $B(\ln z^*, \delta)$ now implies that if $\ln z^* \geq \ln z^{*\prime}$, then

$$\left(\frac{\Pi}{R}\right)_{\ln z^*} \leq \left(\frac{\Pi}{R}\right)_{\ln z^{*\prime}}.$$

If λ is log-linear, then the previous analysis immediately implies that the share of aggregate profits is the same in tougher markets. Since the labor share is $S = 1 - \frac{\Pi}{R}$, we therefore have the following proposition.

Proposition 3 *In tougher markets, the aggregate share of labor in revenues is lower (and the share of aggregate profits higher) if λ is log-convex, the share is higher if λ is log-concave and the share is the same if λ is log-linear*

Appendix A.4 Discussion

Proposition 1 of the model delivers the intuitive result that markups are higher for more productive firms. Thus, the labor share is lower for larger firms. An increase in market toughness that reallocates more output to these firms which will tend to reduce the aggregate labor share. However, a change in market toughness will also change the level of each individual firm's labor share. Greater toughness will tend to increase the elasticity of demand and (from equation 13) push down all individual firm mark-ups and so *increase* the firm-level labor share (a "within firm" effect). Propositions 2 and 3 show that when the underlying productivity distribution is log convex, the reallocation effect dominates the within firm effect so that the aggregate labor share unambiguously falls even though individual firms' labor shares rise. Thus a rise in the aggregate markup does not necessarily indicate a fall in competition—it can mean the opposite.

Proposition 3 also shows that the net effect on the aggregate labor share is an empirical issue: it depends on the shape of the underlying productivity distribution. Interestingly, the standard assumption that the the underlying productivity distribution has a Pareto shape corresponds to a knife-edge case: Pareto is log-linear, and so it produces the result that the aggregate labor share is invariant to changes in market toughness. This is the result in the second part of Melitz and Ottaviano (2008) where they show that the profit share is invariant to changes in market size (L) and competition (γ). Although our proof uses a more general class of demand systems than theirs, we have shown that the reason for their invariance result is due to the assumption of a Pareto distribution for productivity.

Finally, note that the comparative statics on competition abstracts away from entry. If we endogenized entry, there may be a change in the number of entrants and thus in the total expenditure on the sunk cost, κ. What effect this will have on the labor share will partly depend on how this sunk cost breaks down between labor and other factors of production that we have ignored in this Appendix. For example, consider the model of Section II where there are two productive factors, labor and capital. In this case, if the sunk cost is mainly capital and more firms choose to pay the sunk cost to take a productivity draw to enter the more "winner takes all" market, there will be a further fall in the labor share when market toughness rises. If the sunk cost splits in other ways, this is less clear.[58]

[58] A similar issue arises if we close the model and consider how the profits from market power are distributed.

Appendix B MARKUPS

Appendix B.1 Methodology

As noted in the main text, we implement an accounting approach and an econometric approach to estimate markups of price over marginal costs based on equation (7): $m_{it} = \left(\frac{\alpha^v_{it}}{S^v_{it}}\right)$. There are many well-known challenges in performing econometric estimation of production functions, and we apply a variety of approaches to ensure that our conclusions are robust. For our benchmark specification, we follow Section II in estimating a Cobb-Douglas production function separately for each two digit SIC manufacturing industry k:

$$\ln Y_{it} = \alpha^v_k \ln X^v_{it} + \beta_k \ln K_{it} + \ln \theta_{it} + \varepsilon_{it} \tag{23}$$

where $\ln \theta_{it}$ is an unobserved productivity shock and ε_{it} is the unanticipated shock to output (or measurement error). In order to estimate α^v_t, we follow the literature by using a control function approach while modeling $\ln \theta_{it}$ as a first order Markov process. By inverting an input demand equation, we can write productivity as $\ln \theta_{it} = h_{kt}(d_{it}, \ln K_{it})$ where d_{it} could be a dynamic control such as investment (as in Olley and Pakes, 1996) or a static control such as intermediate inputs (as in Levinsohn and Petrin, 2003). Both approaches have two stages. In the first stage, we non-parametrically project output on inputs and the control variable:

$$\ln Y = \phi(\ln X^v_{it}, \ln K_{it}, d_{it}) + \varepsilon_{it} \tag{24}$$

where $\phi_{it} = \alpha^v_k \ln X^v_{it} + \beta_k \ln K_{it} + h_{kt}(d_{it}, \ln K_{it})$. Assuming the productivity process can be written $\ln \theta = g(\ln \theta_{it-1}) + \xi_{it}$ gives rise to the moment condition $E[\xi_{it}(\alpha^v_{kt}) \ln X^v_{it}] = 0$, which can be used to recover the output elasticity. In the second stage we estimate productivity from $\ln \theta_{it} = \phi_{it} - \alpha^v_k \ln X^v_{it} - \beta_k \ln K_{it}$, where ϕ_{it} is recovered from the first stage equation. We can then obtain $\xi_{it}(\alpha^v_{kt})$ by projecting current productivity ($\ln \theta_{it}$) on its lag ($\ln \theta_{it-1}$). The key assumptions underlying this approach are that (1) the variable input responds to productivity shocks but its lag does not; and (2) the lagged variable inputs are correlated with current use of variable inputs (via the persistence in productivity).

A key practical data challenge for both the accounting and econometric approaches to estimating markups is that outside manufacturing we do not observe capital or materials in the Census data. Consequently, in what follows, we perform estimates for manufacturing only. We estimate the production functions at the plant level and then use value-added to aggregate either to the firm level or industry level.

Finally, note that as is standard we observe the nominal values of outputs and inputs, not their quantities. If there is within firm heterogeneity in output prices and input prices then the error term of equation (23) will include the deviations of these firm-specific output and input prices from their industry averages. This will, in general, cause biases to the estimates of the production function parameters and therefore the markups, but it is unclear in which direction. There is now a large literature on how to deal with this issue. First best would be direct observation of the firm specific prices. But while firm-specific output prices are sometimes observed for specific industries and countries (e.g. Foster, Haltiwanger and Syverson, 2008), firm-specific input prices are almost never available (except for imported goods). De Loecker et al (2016) suggest a method for dealing with this noting that the unobservably persistent error in (23) is not just productivity, but actually

It seems reasonable that this is mainly distributed to equity holders, but in principle it could be appropriated by workers in the form of remuneration.

the log of the markup itself, i.e. $\ln p_{it} + \ln \theta_{it} - lnW^v_{it}$.[59] They argue for using a control function to proxy this unobserved variable where the markup is a function of the firm's market share. As discussed in the main paper, they implement this approach in de Loecker, Eeckhout and Unger (2019) for US Compustat data to estimate markups. Like us, they find robust evidence of rising aggregate markups driven by reallocation and emphasis that this is unaffected by the precise way one estimates production functions either using the de Loecker et al (2016) approach or the more standard methods we implement in out paper. Indeed, they find that replacing their estimate of the output elasticity with respect to all variable factors, α^v_{it}, with a constant value of 0.85 leads to near identical results as more sophisticated methods. This is consistent with what we find - the increase in aggregate markups in the Census data is robust to multiple ways of measuring α^v_k. Fundamentally, this is because the increase in aggregate markups is driven by changes in the share of variable input costs in total revenue and not by details of how the output elasticities are estimated in the production function (i.e. the denominator of $m_{it} = \left(\frac{\alpha^v_{it}}{S^v_{it}}\right)$). In our context, we focus on the fall of the labor share as the key variable cost which seems responsible for the estimate of increasing markups.

Appendix B.2 Results

Figure A.2 shows the relationship between firm-level estimated TFP and size. We aggregate the plant-level estimates of TFP using value added shares and use $\ln(\text{sales})$ as a size measure. The ordering of the panels follows those in Figure 10 in the main text. The underlying coefficients to calculate TFP in Panel A uses the accounting method of equation 8. Panel B uses the Levinsohn and Petrin (2003) method of estimating a Cobb-Douglas production function. Panel C does the same as Panel B, but uses the Ackerberg, Caves and Frazer (2015) method of estimating a Cobb-Douglas. Panel D continues using the Ackerberg, Caves and Frazer (2015) method but generalizes Panel C to estimate a translog production function.[60] It is clear that there is a strong positive relationship between size and TFP regardless of the precise way in which the production function is estimated. This is unsurprising as a number of papers have found that productivity and size co-vary positively. Indeed, even using labor productivity we see a similar positive relationship. This is illustrated in Figure A.3, which present the relationship between labor productivity as measured by sales per worker and firm size for each of the six sectors. Recall that the absence of data on intermediate inputs in the Census means we cannot calculate TFP for these sectors. In all six sectors, there is a clear and strong positive relationship between productivity and size. Finally, Figure A.4 shows that large firms have higher markups, as noted in the main text.

Figure 10, which we discussed in the text, reports the results for the baseline accounting and three alternative econometric approaches. The key result is that the aggregate markup has risen substantially, which is of course the flip side of the fall in the labor share. Importantly, the typical firm (i.e., the median or unweighted average firm) has not had a large increase in the markup, whereas the markup at the weighted (by value-added) mean firm increased considerably. This is also consistent with our decomposition analysis.

We have implemented many robustness tests of these findings. First, note that apart from Hicks neutral technical change, we have assumed the production function parameters are stable over time within an industry. However, biased technological change may cause the output elasticities to change over time. To allow for this, we split the sample into two equal time periods (1982-1997 and 1997-2012) and estimated the production function separately in each. We find that the coefficients

[59] More precisely, the firm-specific deviation of these objects from their industry-year averages. Note that marginal costs are variable factor prices adjusted for productivity, $lnW^v_{it} - \ln \theta_{it}$.

[60] Table A.4 indicates our underlying estimates for the output elasticities of the production functions.

are broadly stable over time and the estimated markup trends change little. This calculation is also useful as the fall in the labor share might in theory have been caused by a fall in the output elasticity of labor (see equation 1). Empirically, however, there is no sign of such a decline; in fact, the mean estimated α^L_{kt} across industries rose slightly in the second period relative to the first. In two further robustness tests, we implemented a control function for sample selection following Olley and Pakes (1996), and we included time dummies instead of a time trend in our baseline specifications. Across all of these permutations, we obtained little change to the results.

We also examined how quantiles of the markup have changed over time. Although there has been some increase in the variance of the markup, the changes are not very large. There is some evidence of falling markups in much of the distribution except for the upper tail across all of our estimation methods.

Appendix B.3 Summary

If the output elasticity of labor is constant over time and across firms, then the change in the labor share is the inverse of changes in the markup from equation 1. In this Appendix, we have relaxed this assumption and estimated α^L_{kt} using an accounting method (following Antras et al, 2017) and a production function approach (following de Loecker, Eeckhout and Unger, 2019). We find evidence that complements our main results for the falling labor share. Large firms have higher markups, and aggregate markups have risen in manufacturing. This is primarily due to changes at the right tail of the firm size distribution, with a growing share of sales and value-added accruing to large, high-markup firms.

Appendix C CHARACTERISTICS OF SUPERSTAR FIRMS

We provide additional descriptive evidence on what we term superstar firms based on Standard & Poor's Compustat database. Compustat derives its information from public filings of stock market-listed companies and is thus not subject to the non-disclosure rules that govern our main data from the Economic Census. We focus on the largest 500 firms in Compustat rather than all publicly listed firms, as the population of listed firms has changed substantially and non-randomly over time (see Comin and Philippon, 2006; Davis and Haltiwanger, 2007). The resulting sample will be close to the full set of largest non-government owned companies in the U.S., and thus seems suitable for the analysis of "superstar firms."[61] We focus on the largest firms (top 25, 50 and 500) as defined by sales, but similar results arise if we select the largest firms by employment or market value. All dollar values are inflated to 2015 using the GDP deflator.

Appendix C.1 The 25 Largest U.S. Firms

Table A.5 lists the 25 largest U.S. firms by global sales in 1985, 2000 and 2015. In 1985, the top 25 firms combined for \$1.672 trillion in sales. By 2015, a new set of top 25 firms accounted for sales that were about twice as large in real terms (\$3.748 trillion). There is considerable churning among the top firms, with only General Motors, Ford, Exxon Mobil, Chevron, and AT&T making the top 25 list in each of the three indicated years. Of the top three firms in 2015, only Exxon Mobil's predecessors Exxon and Mobil were already giants in 1985. Walmart, the largest firm in 2015, was just a regional power in 1985, and Apple, the third-largest firm in 2015, was only in its ninth year of operation and still more than two decades away from launching the iconic iPhone in

[61] Compustat includes only firms that have a listing at a U.S. stock exchange, and is thus most complete for firms that are incorporated in the U.S.

2007. Table A.5 also indicates notable changes in industry composition among the largest firms. In 1985, 14 out of the top 25 firms were industrial conglomerates or companies engaged in heavy manufacturing or oil and gas. The representation of these sectors in the top 25 subsequently fell to nine firms in 2000, and to six firms in 2015. Simultaneously, retail, the most rapidly concentrating sector according to our analysis of Census data (see Figure 4), increased its top 25 representation from four to six firms, with Walmart rising to the very top of the ranking. Six of the companies that entered the ranking during the thirty-year window conduct activities associated with healthcare (i.e., pharmacies, drug wholesalers, and health insurance), while four new superstar firms operate in IT-related areas (computer hardware, software, and internet sales). We also see the rise and fall of finance: only one of the top 25 was in banking in 1985 (Citicorp). This number rose to five by 2000 then fell to two in by 2015 (JP Morgan and Bank of America).

Appendix C.2 Growing Firm Size

Figure A.12 provides additional evidence on the evolving size of the 500 largest U.S. firms, which increased strongly over the last four decades. In 1972, the combined global sales of the 500 largest U.S. firms was about \$3 trillion. By 2015, this value was nearly \$12 trillion. Market value expanded even faster, by a factor of six rather than three.[62] Employment in the top 500 firms grew at a considerably slower pace, however, increasing by only 60 percent.

This growth does not merely reflect the overall expansion of the U.S. economy. Panel A in Figure A.13 plots the ratio of top 500 firms' sales to gross output of the U.S. private sector. This ratio increased from about 0.37 in 1972 to 0.40 in 2015. The increase was not monotone through this time period, however, but rather featured substantial fluctuations. Some of these fluctuations, especially during the 1970s and 1980s, were due to large changes in oil prices that translated to a high volatility in oil firms' sales over time. The real oil price more than doubled from 1973 to 1981, and the oil industry accounted for seven of the eleven largest U.S. firms by sales in 1981. After 1981, the oil price declined rapidly, and with it the sales of oil firms. To purge variation that stems from the cyclical behavior of oil prices, Panel B in Figure A.13 indicates the ratio of top 500 non-oil firms' sales to the gross output of the U.S. private non-oil sector. This time series shows a stronger and less volatile growth of the largest firms' sales relative to U.S. output. The pattern of relatively rapid sales growth in the largest U.S. firms is consistent with the overall increase in concentration documented in the Economic Census data throughout this paper.

Appendix C.3 Inequality Among the Largest firms

We next investigate whether concentration has risen *among* the top 500 superstar firms. Figure A.14 plots the share of the largest 50 firms in the combined sales of the largest 500 firms. Over the full period, this share grew from 42 to 48 percent. By the end of the sample period, the largest 50 firms thus accounted for almost the same volume of sales as the next largest 450 firms combined. However, unlike the growth in size of the top 500 firms (Figure A.12), the growth of concentration among these largest firms was not rising monotonically over time. Instead, Figure A.14 shows that sales concentration was weakly falling until the late 1990s then increased until 2010 and leveled off thereafter.

[62] The market value reported in Figure A.12 corresponds to the numerator of Tobin's Q, as in Gabaix and Landier (2008). It is computed by summing up the stock market value (number of shares outstanding times closing stock price) and the value of debt (long-term debt and current liabilities). We obtain a similar time series for stock market value alone.

Figure A.15 provides additional evidence for the rising concentration in sales among the largest 500 firms by examining changes in the cross sectional dispersion of sales among these firms. The first panel plots the time series for the mean, median, and 5th and 95th percentile of sales among the top 500 largest firms. The quantiles of the size distribution have fanned out over time, with the growth in the upper tail of the distribution (e.g. between firms at the 95th percentile and the median) being particularly stark. Sales growth has been stronger for the mean than the median and for the upper quantiles compared to the mean. The second panel normalizes each series to one at the start of the period and shows that the *relative* level of sales has become considerably more dispersed among the top 500 firms since about the year 2000.

The fact that the growth of sales concentration among large firms increases only after 2000 may come as a surprise given that we can see concentration rising since 1982 in the Census data. Several factors may contribute to this pattern. First, publicly listed firms account for only a fraction of all economic activity in the U.S. whereas the Census covers all firms in a given sector. Second, most sales and employment data in Compustat relates to globally consolidated accounts covering both firms' operations in the U.S. and abroad, which is distinct from the Census' exclusive coverage of domestic U.S. employees and sales by domestic establishments. The most important reason, however, is that our Census analysis focuses on concentration within four-digit industries whereas the Compustat analysis combines firms from all sectors because it comprises a much lower number of firms. When we perform an analogous exercise in the Census data, ignoring industry, we obtain a time series of concentration much more similar to the one found in Compustat.

Appendix C.4 Dynamics

Growing concentration could be consistent with greater churn among the the largest 500 firms ("creative destruction") or decreasing churn ("persistent dominance"). Figure A.16 shows the fraction of the top 500 sales firms in each year that were among the 500 largest firms one, five, and ten years previously. It indicates that churning among the largest firms (at least for five and ten year churn rates) rose in the pre-2000 period, but has fallen since 2000—the period where we have shown that concentration rose. For example, of the firms that comprised the top 500 in the year 2000, two-thirds were already in the top 500 five years earlier. By the end of our sample period, the five-year survival rate in the top 500 had risen to more than eighty percent. Census data also show declining churn since the 2000 period (see Decker, Haltiwanger, Jarmin and Miranda, 2018). So increasing inequality between firms seems to be accompanied by more persistent dominance rather than greater creative destruction.

Appendix C.5 Activity across Countries and Industries

One possible explanation for the rapidly growing size of superstar firms is the increasingly global scale of their operations. While the Compustat data reported above correspond to firms' worldwide activities, most U.S.-based Compustat firms also report a breakdown of their revenue between domestic and international sales since 1978.[63] Figure A.17 documents that the 500 largest U.S. firms on average sold around about 20 percent of their output in foreign markets in the early 1980s.

[63] A breakdown of the top 500 firms' domestic versus foreign sales based on geographic segment data is missing for 13% of all firm-years. We sequentially impute missing foreign sales shares by (i) using data on the geographic composition of sales from operating segment data, (ii) using the the foreign sales share of the nearest year in which the firm did report this information, or linearly interpolating if the geographic breakdown is available for both an earlier and a later year, and (iii) imputing the foreign sales share using the average value for the same year of other top 500 firms of the same 2-digit industry, or of the same broad sector (manufacturing or non-manufacturing) if there are no other firms from the same industry. Compustat does usually not report a geographic breakdown of employment.

Foreign sales grew rapidly in importance during the 1990s and the 2000s, and accounted for more than 35 percent of the sales of top 500 firms by 2010. The growth in foreign sales during the 1990s and 2000s coincides not only with a rapid expansion of international trade but also with greater foreign direct investment. For instance, Walmart has exported its successful business model to several countries in Latin America, Europe and Asia, and now generates nearly 30 percent of its total sales abroad according to the Compustat data.

Another potential source of superstar firms' growth is an expansion of activity across industries. Berkshire Hathaway, one of the five largest U.S. companies by sales in 2015, operates across an eclectic range of industries from insurance to confectionery, railroads, home furnishing, newspapers, and energy. The retail giant Amazon also has extended its reach into a large number of different markets. We explored in the Census data whether firms that are among the top four sellers in a four-digit industry have increasingly become dominant players across in other four-digit industries as well. We do not find that there is a general trend towards greater diversification across industries among firms. The largest firm (by sales) in the four-digit industry in the Census operated on average in over 13 other four-digit industries in 1982, but this number fell to under nine by 2012. Similarly, conditional on a firm being among the top four firms (again by sales) in a four-digit industry in 1982, it was among the top four in 0.37 other industries in that same year (i.e. statistically speaking, being the top firm in one industry gave a firm almost a 40 percent chance of being among the top four in one other industry). This fraction fell to 0.24 by 2012. Thus, the "Amazon" pattern, where one firm appears to become dominant in multiple industries, does not seem to be representative of what is occurring among the largest firms.

Appendix C.6 Labor Share Trends in Compustat

In addition to the limitations imposed by partial coverage and the aggregation of firms' U.S. and global activities, the Compustat data presents several additional data issues for analyzing labor shares. First, labor costs are not a mandatory reporting item for publicly listed U.S. firms—only about 13 percent of firms report "staff expenses," and those reporting are mainly larger firms. Second, value-added is not reported in Compustat as there is no consistent definition of intermediate inputs.

Despite these multiple caveats, we obtain broadly similar patterns of results for labor share when examining Compustat data. For purposes of the Compustat analysis, we define the labor share as the ratio of wage bill to the best proxy for value added–the sum of wage bill and EBITDA (earnings before interest, tax, depreciation and amortization). There is a clear decline in the aggregate labor share from nearly 60 percent in the early 1980s to 47 percent in 2015 in the subset of top 500 firms for which data are available. Figure A.18 reports the change in labor share separately for firms whose share of foreign sales in total sales is above or below the median of a firm's 2-digit industry. Firms with more global engagement have higher labor shares on average in most years. While the data usually do not provide a breakdown of employment by location, it is possible that firms with greater foreign sales also have a larger proportion of their production employment abroad, and major economies like Germany or the United Kingdom have higher labor shares than the U.S. as shown in Figure 1. Despite this difference in the levels of labor shares, firms with greater or smaller foreign sales shares both experience very similar declines in the labor share over the full 1978 to 2015 period. The commonality of labor share declines among more and less globally engaged firms suggests that, while globalization may be one factor behind the trend of declining labor shares, it is unlikely to be the whole story.

Appendix D DATA

Appendix D.1 Economic Census

Our primary data are from the U.S. Economic Census conducted every five years by the Census Bureau.[64] We focus on six sectors for which we could access micro-data over a significant period of time: manufacturing, retail trade, wholesale trade, services, utilities and transportation and finance. There is also a Census of Construction, but it does not provide a consistent firm identifier. Within these six sectors, several industries are excluded from the Economic Census: rail transportation from transportation; postal service from wholesale trade; funds, trusts and other financial vehicles are excluded from finance; and schools (elementary, secondary, and colleges), religious organizations, political organizations, labor unions and private households are excluded from services. The Economic Census also does not cover government-owned establishments within covered industries.

Our analysis includes only establishments that have at least one employee ("employer firms"), a positive value of annual sales, value-added, assets, material costs, and salaries and wages, and are assigned a code that allows us to link them over time in the Census (LBDNUM). We exclude any observations that are drawn from administrative records, as these observations are largely imputed and are not included in official statistics published by the Census Bureau. We also winsorize the establishment-level labor share at the 99^{th} percentile to account for outliers. As an establishment's value-added goes towards zero, the labor share can become arbitrarily large. While this has little effect on the industry-level analysis, where we weight observations by their share of value-added, these large outliers can affect the decomposition of changes in labor share into between-firm reallocation and within-firm components in Figure 7 and 8. We confirmed the robustness of our results to alternative treatments of outliers, including dropping them altogether or top-coding the labor share at one.

While each establishment is assigned to one primary industry, firms with multiple establishments are often active in several industries. In all of our industry-level analyses, we define firms separately by four-digit SIC industry, meaning that a firm with establishments in three different industries will be treated as three separate firms in our analysis. This definition of the firm is motivated by our focus on concentration ratios, where the relevant measure is not the total size of the firm but rather the importance of that firm in a given industry. In manufacturing, about 20 percent of firms are active in multiple industries, and on average, firms span 2.6 industries. These numbers are slightly lower in retail and wholesale trade and services, but are slightly higher in finance where about a quarter of firms span multiple industries. The only analysis in which we do not define a firm as a firm-by-industry pair is the overall within-between decomposition in Table 5 and 4. In this table, we define a firm using all establishments that belong to the same broad sector and are thus covered in the same segment of the Economic Census. However, in Appendix Table A.1, we present decomposition in which we define a firm using the firm-by-industry pair.

The sales measure in Census is shipments so it includes exports. Since the labor used at the firm goes into the production of output destined for exports as well as domestic consumption, it seems natural to use total sales. The concentration measures published by the U.S. Census Bureau also follow this convention. If we wanted a purely domestic measure of market concentration, we would want to deduct exports in concentration measures.

[64] More details on Economic Census are available at https://factfinder.census.gov/faces/affhelp/jsf/pages/metadata.xhtml?lang=en&type=survey&id=survey.en.ECN_ECN_US

Appendix D.2 Constructing Time-Consistent Industry Codes

Since we analyze cross-industry variation in concentration, accurate classification of industries is central to our analysis. In the raw data, each establishment is assigned an industry code that is based on the primary activity of the establishment. In 1982, the establishments are given a 1972 SIC code, from 1987 to 1997, the establishments are given a 1987 SIC code, and from 1997 to 2012, the establishments are given a NAICS code based on the classification corresponding to that year (i.e. 2002 is in 2002 NAICS codes). While most of our regressions are run at the industry level, the definition of industry concentration ratios and firm-level decompositions requires that each establishment is assigned to a single industry, meaning that a weighed (i.e., fractional) crosswalk of NAICS to SIC codes is not suitable. To construct a one-to-one crosswalk, we utilize the panel structure of the Census data and the fact that in 1997, each establishment is given both a 1987 SIC code and a 1997 NAICS code. If the establishment has the same NAICS code in the following years, we assign the given 1987 SIC code that is reported for the year 1997 to the later years as well. Then, if either the establishment was not in the sample in 1997 or the NAICS code changed in the later years, we use a modal mapping from the NAICS codes to the 1987 SIC code, meaning that we assign each NAICS industry to the SIC code that is it most likely to map to in the probabilistic mappings provided by the Census.

There are, however, some 1987 SIC codes that are not the most likely industry for any NAICS code, meaning that those 1987 SIC industries would not exist in the post-1997 data ("orphaned SIC codes"). To avoid the creation of such an artifact in the data, we aggregate SIC codes so that each aggregate SIC code is observed both before and after the SIC-NAICS seam. In deciding which industries to group, we find the 1997 NAICS codes that establishments from the orphaned SIC codes are most likely to be reclassified as, and then we combine that SIC code with the SIC codes that were the most likely 1987 SIC codes for that NAICS code. For example, establishments from 1987 SIC code 2259 "Knitting Mills, Not Elsewhere Classified" are most likely to be re-classified as NAICS code 315191 "Outerwear Knitting Mills." But of all the establishments that were given code 315191, the most common 1987 SIC code was 2253 "Knit Outerwear Mills." Therefore, we aggregate the 1987 SIC codes 2253 and 2259. We follow the same procedure for bridging the 1972-1987 SIC reclassification.

Finally, we were forced to exclude some industries that are not defined consistently over time in the Census. These are only in manufacturing, services and finance. From manufacturing, we drop industries the move outside manufacturing in the 1997 SIC-NAICS redefinition which are 2411 (Logging), 2711 (Newspaper Publishing and Printing), 2721 (Periodical Publishing and Printing), 2731 (Book Publishing and Printing), 2741 (Miscellaneous Publishing), 2771 (Greeting Cards) and 3732 (Boat Building and Repair). From Services, we drop SIC codes 7338 (Secretarial and Court Reporting Services), 8734 (Testing Laboratories), 8062 (General Medical and Surgical Hospitals), 8063 (Psychiatric Hospitals), and 8069 (Specialty Hospitals, Except Psychiatric). From Finance, we drop SIC codes 6722 (Management Investment Offices), 6726 (Unit Investment Trusts), 6552 (Land Subdividers and Developers), 6712 (Offices of Bank Holding Companies) and 6719 (Offices of Holding Companies not elsewhere classified).

Our final industry panel corresponds to a slight aggregation of four-digit SIC industries, and comprises 388 industries in manufacturing, 58 industries in retail trade, 95 industries in services, 31 industries in finance, 56 industries in wholesale trade, and 48 industries in utilities and transportation.

There are, of course, other ways of constructing consistent industry codes in the Census. A leading alternative is Fort and Klimek (2016), detailed in their Appendix A.They use NAICS codes based in 2002 to code every LBD establishment. They do this by first using longitudinal data in LBD to fill in missing codes then they use concordances to assign all NAICS codes that map uniquely to a SIC code (i.e. NAICS codes that are full contained in a SIC code). They next use the longitudinal structure to assign NAICS codes to an establishment with an SIC code that maps to many NAICS. Finally, in instances where the longitudinal information is insufficient and the SIC code maps to multiple NAICS codes, they use random assignment to assign a NAICS code. In order to do a robustness check, we restricted our analysis to the set of six-digit NAICS industry codes that are consistently reported over time. These cover close to 98 percent of all employment and sales in our six Census sectors. We then calculate concentration and labor shares for this subset of industries and re-ran the analysis. We find results that are very similar to the main ones we report in the paper. For example, in the first three columns of Table 3 (the labor share vs. concentration regressions), all 18 coefficients across the six segments are negative and 16 of these coefficients are significant at the 5 percent level or greater.

Appendix D.3 Correcting Census Value-Added for Service Intermediate Inputs using KLEMS

The measure of value-added in the Census of Manufactures adjusts for intermediate purchased goods. It is defined as sales (item TVS) less inventory investment for final goods (difference between FIE and FIB) and work in progress goods (difference between WIE and WIB), resales (item CR), material inputs (sum of items CP, CW and MIB less MIE) and energy expenditures (sum of items CF and EE). This definition does not adjust for intermediate purchased services, however,[65] meaning that an increase over time in intermediate purchased services will appear in the Census data as an increase in value-added (and possibly exaggerates the fall in the labor share). The KLEMS data provide information on use of intermediate services by U.S. industries based on the input-output framework of the BEA. They thus allow us to roughly adjust value-added in the Census to account for any trends in intermediate purchased services over time. Since the KLEMS data are only available at the two- to three-digit industry level, we make the adjustment at the establishment level in two ways, both of which use the fact that the Census data include information on the value of material costs for each establishment. First, we calculate in KLEMS the ratio of intermediate purchased services to intermediate materials and assume that each establishment in a given two-digit industry utilizes purchased services in that proportion. This is the method we report in Row 3 in Table 2. As a second alternative, we calculate the fraction of total two-digit industry intermediate material costs that are accounted for by each four-digit industry, and assume that four-digit industries purchase the same fraction of total intermediate services. The level of the labor share is higher (as value-added is lower) with either correction for purchases of intermediate services, but the trends are similar across the original and adjusted data series.

Appendix D.4 Comparing Census and NIPA/BEA data

In this subsection, we compare the Census data that we use throughout the analysis to the broad industry-level NIPA data produced by the Bureau of Economic Analysis (which is used by Elsby, Hobjin and Sahin, 2013, for example). The goal of this exercise is twofold. First, we aim to validate the construction of establishment-level data by showing that, when aggregated, it is similar to the aggregate trends discussed widely in the literature. Second, we use the NIPA data to benchmark the payroll-to-sales ratio outside of manufacturing to Census data. Since the Census does not collect sufficient information outside manufacturing to construct measures of value-added, our main analysis uses the payroll-to-sales ratio as an alternate measure.

[65] There is a deduction for contract work, CW, but this is narrowly defined.

The Census derives its estimates from mandatory report forms. The NIPA estimates are instead derived from a compilation of data sources. One of these sources is the Economic Census, but it also includes annual, quarterly and monthly surveys, financial reports, government budgets and IRS tax data. A reason for these additional data is that NIPA data are reported at a higher frequency (quarterly) than Census data. They are also reported at a higher level of industry aggregation than Census. For our purposes, this difference leads to two important distinctions between the Census and NIPA data. First, the industry definition varies across the two sources. The Census unit of analysis is an establishment whereas in NIPA it is the firm. Consider a firm whose primary industry is retail but that also has a manufacturing plant. In Census data, the employment of the manufacturing establishment is counted towards the manufacturing sector while the remainder of the firm's establishments are classified as retail. By contrast, NIPA could attribute all the firm's employment (including that of the manufacturing establishment) to retail. Additionally, the BEA/NIPA includes some sub-industries that are not included in the Census, such as management and private households.

A second distinction between BEA/NIPA and Census is that the two agencies define the components of the labor share differently. Panel A of Figure A.6 displays the payroll-to-value-added ratio for manufacturing in NIPA and Census, and shows that while the trends are similar, the level of the series differs substantially across the two data sources. As is shown in Panel B of Figure A.6, this discrepancy stems from a small difference in the numerator (compensation) and a larger difference in the denominator (value-added). The first figure in Panel B plots the compensation series in the two datasets, which appear reasonably comparable. As discussed above, there is a narrow and broad definition of payroll in the Census. There is also a narrow and broad definition in NIPA, although the broad NIPA definition is even wider than in the Census. Indeed, the broader definition of compensation in the Census data closely tracks the narrower definition of compensation in the NIPA data.[66]

NIPA and Census data diverge more in their definition of value-added. The second figure in Panel B shows that value-added in the Census data is significantly higher than value-added in the NIPA data. While there are several differences in the two series, the largest difference is in their treatment of intermediate purchased services. Since the Census does not collect information on intermediate purchased services, it does not subtract these from value-added, and therefore measures value-added as the establishment's output less its material costs.[67] However, the BEA does collect information on intermediate purchased services and subtracts it from its value-added measure. To explore the importance of this mechanism, as discussed in the previous subsection we use industry-level estimates of intermediate purchased services from the KLEMS data. These data are reported annually beginning in 1997 at the three-digit NAICS level. As the red line in the right figure of Panel C shows, subtracting off the intermediate purchased services within manufacturing almost exactly closes the gap in value-added across the two data sources. Indeed, using this modified value-added series results in aggregate labor shares from the Census that are near identical to those from NIPA when we use the broader measure of Census compensation (see Panel A of Figure A.6).

As discussed above, the Census does not collect detailed information on intermediate inputs

[66]The BEA also includes a more comprehensive measure of compensation that includes employer contributions to insurance plans as well as government social insurance programs. This is reported on an accrual basis, and reflects liabilities rather than actual payments.

[67]Note that the Census does collect information on the costs of contract work that is done by others on materials furnished by the reporting establishment. Since this cost is included in their measure of intermediate costs, it is subtracted from value-added. However, this does not include the costs of contracted services such as advertising, insurance, or professional consultants.

outside manufacturing. Therefore we analyze the behavior of the payroll-to-sales ratio. Figure A.7 shows the trend of the payroll to value added ratio in NIPA in each of our six sectors. As is well known, there is a clear downwards trend in these series since the 1980s in most sectors.

Figure A.8 shows for each sector the payroll-to-sales ratio in the Census compared with its closest counterpart in NIPA: the payroll to gross output ratio. We also include the NIPA payroll to value-added ratio which is not available in the Census except for the manufacturing sector. Each series is normalized to one in 1987. Starting with manufacturing in the top left panel, the series are relatively aligned in terms of trends, but diverge a bit, especially after 1997. This is mainly because the NIPA data are released in 1987 SIC codes pre-1997 and in 1997 NAICS codes post-1997, creating a discrepancy in the NIPA series.

Looking at the other five sectors, two patterns emerge. First, there is a general downward trend in the labor share measured across almost all sectors. Second, the NIPA trends are more closely correlated with each other than they are with the Census trends, which is unsurprising as the denominator is identical. Third, the Census trends diverge from the NIPA more strongly outside manufacturing, especially around the industry re-classification seam of 1997, which distorts the NIPA series.

Disaggregating the numerator and denominator reveals that the payroll measures in Census and NIPA move much more in tandem than the sales and output measures. Apart from the industry reclassification, there may be several reasons for this divergence in sales and output. First, measuring output in finance poses particular problems as we noted in the main text. In most sectors, BEA uses the Economic Censuses to construct gross output and then they work through data sources on intermediate inputs use to construct value added. For finance, however, BEA uses an entirely different approach using interest rate spreads between lending and deposit rates. This could be a reason for the large discrepancies we see in finance where the labor share falls in NIPA after 1992 but rises in the Census data (at least until 2002). For these reasons, we reiterate that the results for the Finance sector must be treated with the most caution. Second, Census sales differ from NIPA output primarily because of inventories, so output will exceed sales when inventories are rising as a fraction of output. This may particularly be an issue for wholesaling, which will plausibly be strongly affected by inventory behavior, and where we do see large divergences with labor shares rising in the 1987-2002 period in the Census while declining in NIPA. Third, we have excluded some industries that are not defined consistently over time in the Census but are unable to remove these industries from NIPA. So to the extent these sub-industries exhibit different growth trends, this will show up in the aggregates. These dropped industries are exclusively in finance, services and manufacturing.

Appendix D.5 Decomposition Analysis: Details and Robustness

The decomposition analysis is described in the text. In this subsection, we describe some of the robustness tests that we implemented. The baseline analysis treats the firm as the unit of observation, so we aggregate all activity across the establishments belonging to a firm at a point of time in a Census segment. We also confirmed robustness to implementing the decompositions at the establishment level and at the firm by four-digit industry level.

We next considered a generalization of the decomposition breaking out the between industry component. As noted in the text, we first use a standard shift-share technique as in Autor, Katz and Krueger (1998) to decompose the overall change in the labor share into between-industry $\sum_j \left(\widetilde{S}_j \Delta\omega_j\right)$ and within-industry $\sum_j \left(\widetilde{\omega}_j \Delta S_j\right)$ components:

$$\Delta S = \sum_j \left(\widetilde{S}_j \Delta \omega_j \right) + \sum_j \left(\widetilde{\omega}_j \Delta S_j \right). \tag{25}$$

Here, $\widetilde{S}_j$ is the time average of the (size-weighted mean) labor share, S_j, in industry j over the two time periods t_0 and t_1, and $\widetilde{\omega}_j$ is the time average of ω_j , the industry size share (e.g. value-added share of industry j in total manufacturing value added). Thus, the first term in this equation is the change in labor share due to shifts in industry size shares, holding average industry labor shares constant, while the second term is the change in labor share due to within-industry labor share shifts, holding average industry size shares constant. We next re-write our primary Melitz-Polanec decomposition (equation 5) at the industry level:

$$\Delta S_j = \Delta \bar{S}_{S,j} + \Delta \left[\sum_{i \in j} \left(\omega_{i,j} - \bar{\omega_j} \right) \left(S_{i,j} - \bar{S_j} \right) \right]_{S,j} \tag{26}$$

$$+ \sum_{i \in j} \omega_{X,0,i,j} \left(S_{S,0,i,j} - S_{X,0,i,j} \right) + \sum_{i \in j} \omega_{E,1,i,j} \left(S_{E,1,i,j} - S_{S,1,i,j} \right). \tag{27}$$

This notation makes explicit that the labor share of firm i (what we called S_i in equation 5) is also in industry j, so we now denote it explicitly $S_{i,j}$ and similarly for the firm size shares, $\omega_{i,j}$. Substituting equation (26) into equation (25) gives us a decomposition with five terms:

$$\Delta S = \sum_j \left(\widetilde{S}_j \Delta \omega_j \right) + \sum_j \widetilde{\omega}_j \Delta \bar{S}_{S,j} + \sum_j \widetilde{\omega}_j \Delta \left[\sum_{i \in j} \left(\omega_{i,j} - \bar{\omega_j} \right) \left(S_{i,j} - \bar{S_j} \right) \right]_{S,j}$$

$$+ \sum_j \widetilde{\omega}_j \sum_{i \in j} \omega_{X,0,i,j} \left(S_{S,0,i,j} - S_{X,0,i,j} \right) + \sum \widetilde{\omega}_j \sum_{i \in j} \omega_{E,1,i,j} \left(S_{E,1,i,j} - S_{S,1,i,j} \right) \tag{28}$$

A complication arises in equation (28) because we have to determine which four-digit SIC industry a firm (or plant) belongs in. We follow the Census attribution of an establishment to a four-digit industry (based on the amount of shipments in the product trailer). For multi-plant firms that span several industries, we set the main industry as the one which produces the most shipments within the firm. A further complication arises from the fact that plants and firms frequently switch their main industry, especially over the long 30-year period of time we consider (see Bernard, Redding and Schott, 2010). Using a time varying firm-industry definition attributes a large fraction of the changes to entry and exit, even though this type of churn may simply reflect a firm experiencing differential sales growth of one of its products.[68] Hence, in our main implementation of equation (28), we fix the firm's industry to be that in the first year we observe the firm. We also implemented a permutation where we fix the industry designation as the one observed in the last year that the firm is observed in the data (or use the modal industry across all years). These adjustments make no material difference to the results.

Following the discussion of comparing the Census to NIPA labor shares above, we also implemented the baseline decomposition correcting for intermediate inputs using the NIPA. We take the fall in the NIPA labor share (ΔS^{NIPA}) as accurate and then calculate the contribution of each of

[68]For example, consider a plant in period t_0 that ships six units of product A and four units of B and so will be allocated by the Census to industry A. If in period t_1, the units of A stay the same, but it expands shipments of product B to seven, the plant's will now be allocated to industry B. In the decomposition analysis it will be classified as an exit after period t_0 and an entrant in period t_1, whereas in fact it has just shifted its sales portfolio a bit.

the four components (within, reallocation, exit, entry) using the Census decomposition in Table 5. We assume that the fraction of the fall accounted for by each component is the ratio of the Census component to the sum of the (absolute values) of all Census components. Formally, define the contribution of component d as $C_d^{NIPA} = \Delta S^{NIPA} \times \left(\frac{C_d}{\Sigma_{d=1,2,3,4}|C_d|}\right)$ where C_d is the contribution as calculated in Table 5 and $|C_d|$ is the absolute value of this. Figure A.9 shows the results graphically.

Appendix D.6 International Datasets

The KLEMS data derives from an international research collaboration that provides harmonized industry-level information on output, inputs and productivity taken from national statistical agencies. We use the U.S. KLEMS data to measure purchases of intermediate services, as detailed in Section Appendix D.3. We also draw on the 2012 release of the EU KLEMS database (see O'Mahony and Timmer, 2009, http://www.euklems.net/) in order to compare levels and changes in industry-level labor shares across countries.

In addition, we draw on two international firm-level datasets: BVD Orbis and COMPNET. Bureau Van Dijk (BVD) is a private sector aggregator of company accounting data. The panel data set Orbis is its most comprehensive product covering in principle the population of all public and private company accounts in the world (see Kalemli-Özcan et al., 2015; Gopinath et al., 2017). BVD seeks to harmonize the data in a common format focusing on a subset of the variables that are used for investment analysis. Orbis has been built up over time, so it is less comprehensive the further back in time one goes (see Bajgar et al, 2018b). Furthermore, the data are constrained by what firms report in their accounts. Accounting regulations differ across countries with some countries requiring more comprehensive reporting than others. For example, the U.S. requires private firms to report very little information in the public domain compared to European countries such as France. Across all countries, more information is demanded from larger firms than smaller firms.

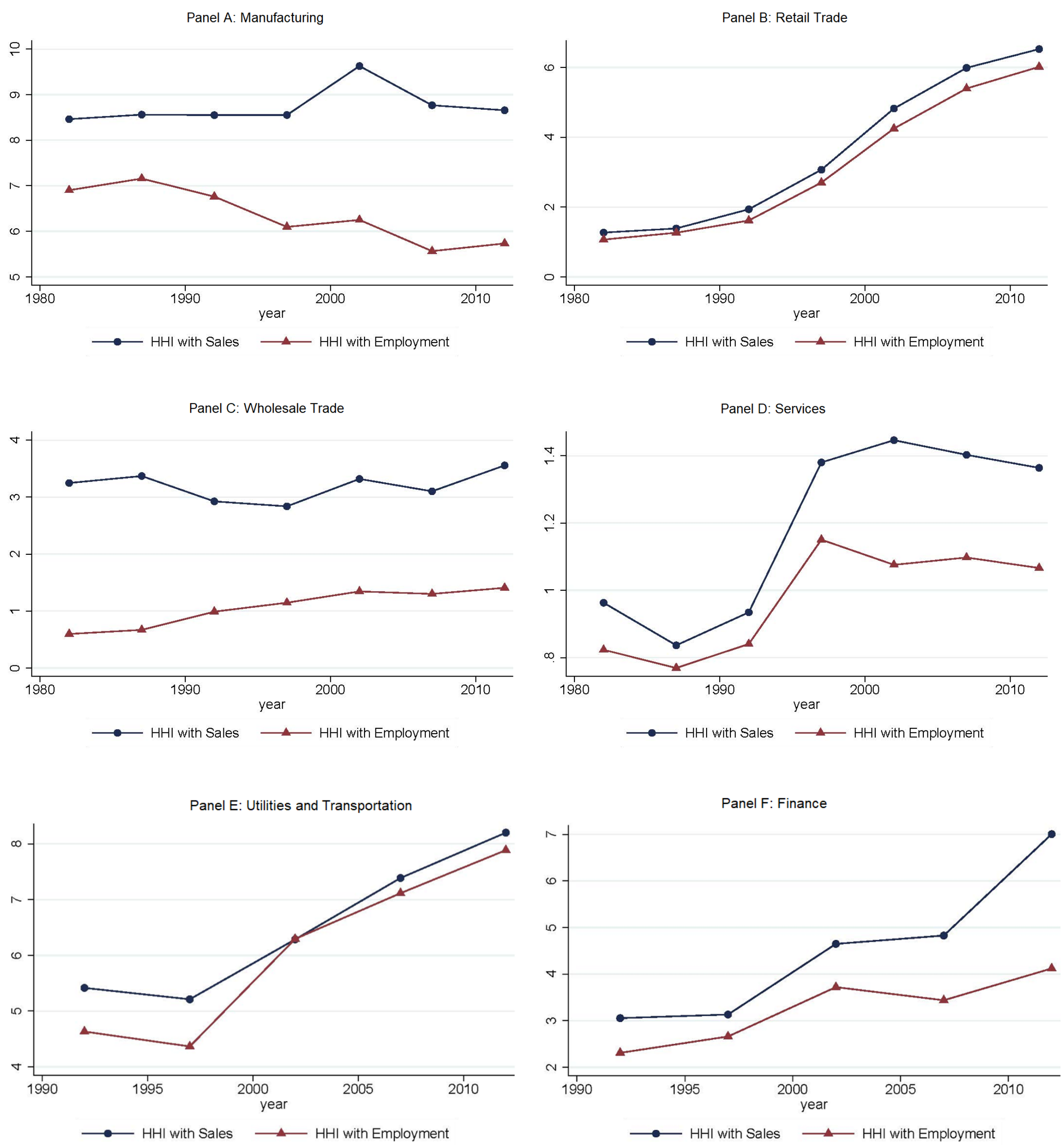
For our analysis we require that firms have information on their primary industry and their payroll. To construct value-added, we sum payroll with gross profits (i.e. before tax, depreciation and interest have been deducted (i.e., EBITA). Intermediate inputs are rarely reported in company accounts, so deducting these from sales (as we do with the Census data) is not feasible. The labor share is then the ratio of payroll to this measure of value-added. We also do some robustness checks comparing this measure with the ratio of wage bill to sales. We focused on the sub-sample of countries where we could get reasonably comprehensive data, and on the five year period with most comprehensive firm coverage for each country, which is 2003 to 2008 for the UK, Sweden and France, and 2005 to 2010 for Germany, Italy and Portugal.

The second international firm database is Compnet. Compnet has balance sheet data from 14 European countries that cover the 2000-2012 period. These data, compiled by the European Central Bank's Competitiveness Research Network, draw on various administrative and public sources across countries, and aim to collect information for all non-financial corporations (see Lopez-Garcia, di Mauro and CompNet Task Force 2015 for details). This was an initiative led by the European Central Bank in a effort to obtain systematic micro-data to help inform its macroeconomic modeling. It was able to coordinate with the Central Banks from different European Union member states to get access to micro-data that were not always in the public domain.

The version of Compnet made available to us (kindly through Erik Bartelsman) aggregates the firm level data to the industry level. It contains information on the labor share and industry concentration (both the fraction of sales produced by the largest ten firms and the Herfindahl-Hirschman Index for various two-digit industries). Although great effort was invested to make these measures comparable across countries, there are some important differences that affect the reliability of cross-country comparisons. Most importantly for our purposes, countries use different reporting thresholds in the definition of their sampling frames. We weight the data to attempt to account for different firm sizes and sample response probabilities.

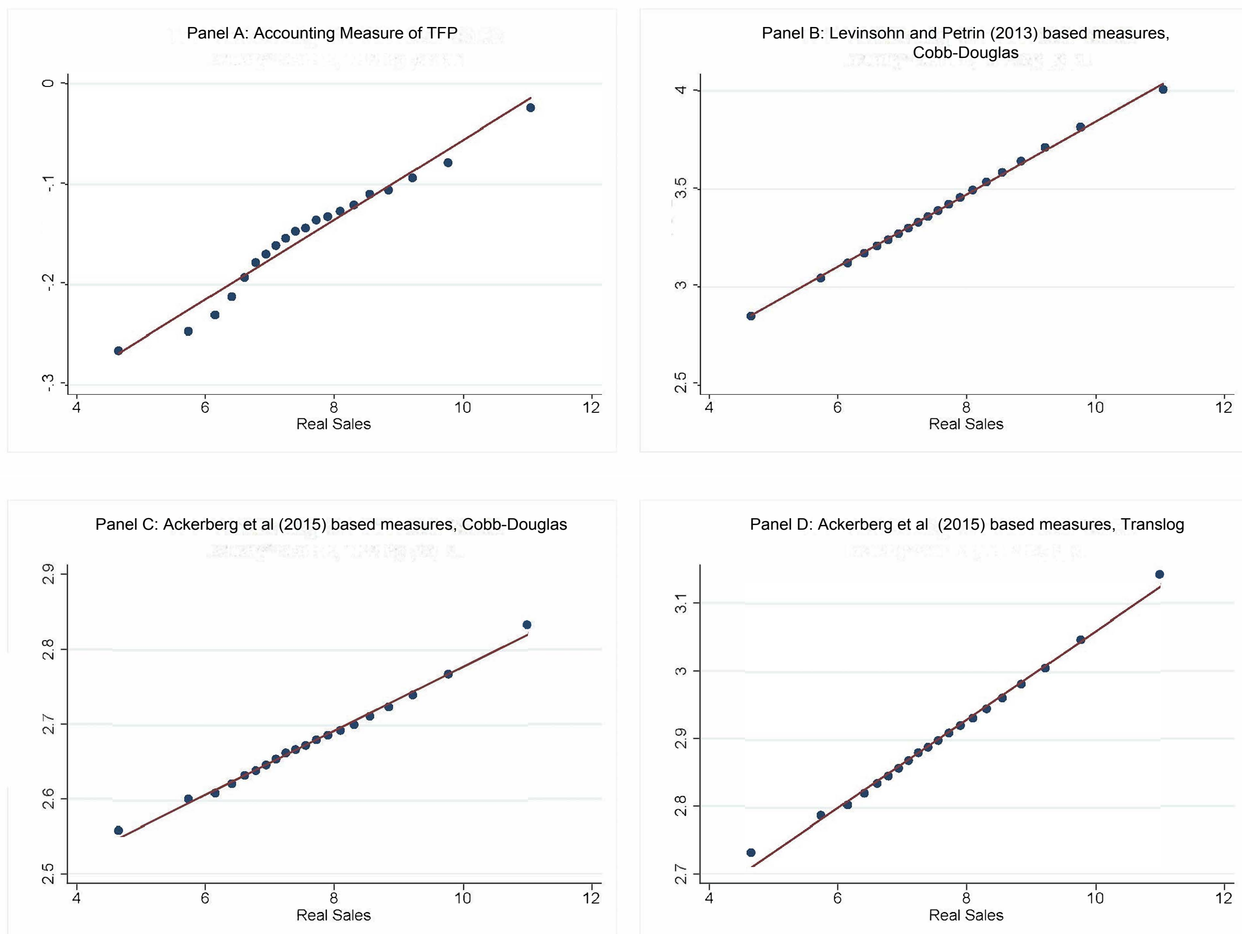
Appendix Figures

Figure A.1: Average Herfindahl-Hirschman Index by Sector



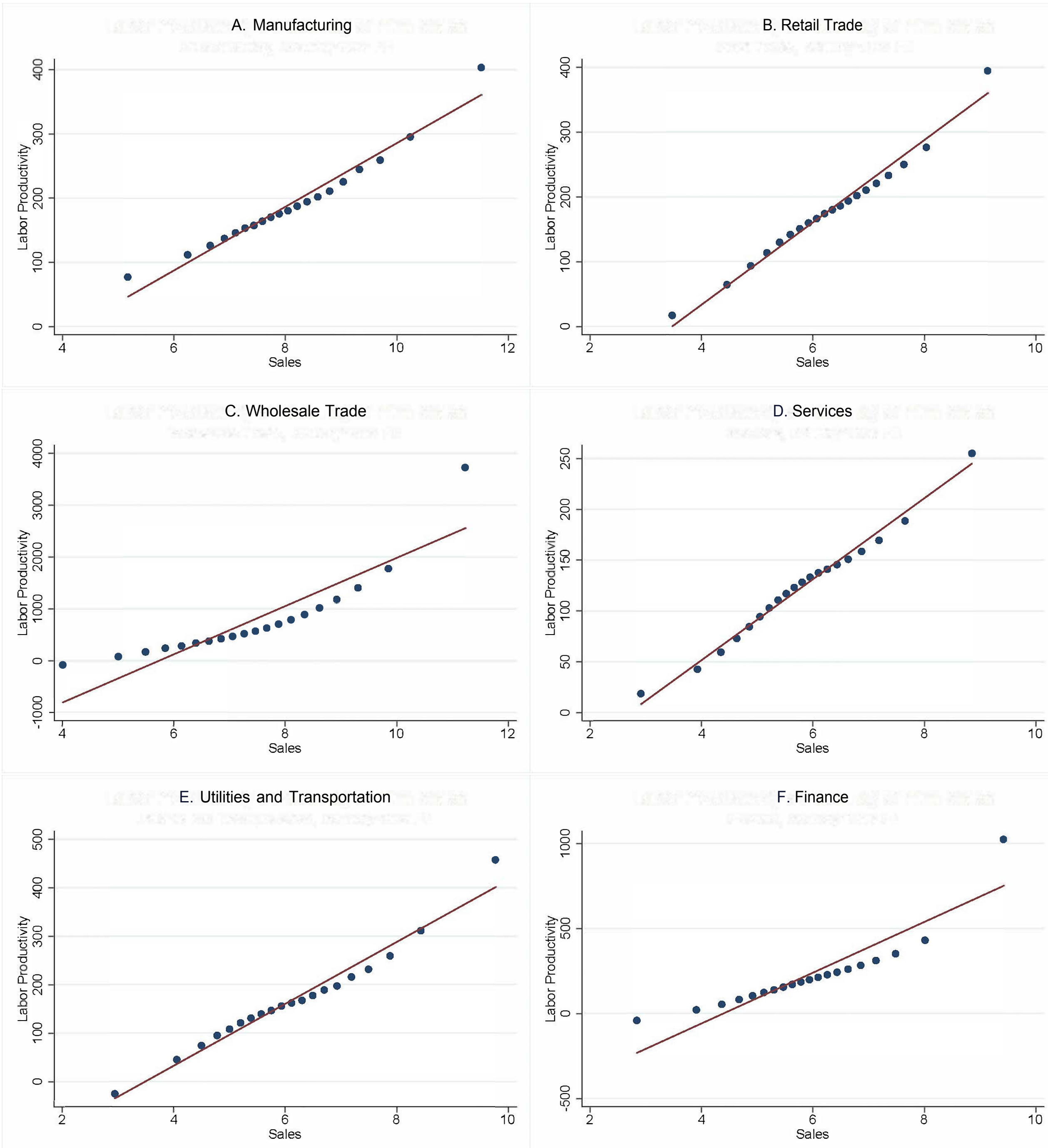
Notes. Each figure plots the average HHI calculated within four-digit industries. Industry concentration is calculated for each time-consistent four-digit industry code, and then averaged across all industries within each of the six sectors using the industry's share of total sales as the weight. The blue circles plot the HHI calculated using firm sales and the red triangles plot the HHI calculated using employment.

Figure A.2: The Relationship between Estimated Total Factor Productivity and Firm Sales Using Four Methods of Estimating TFP, Manufacturing



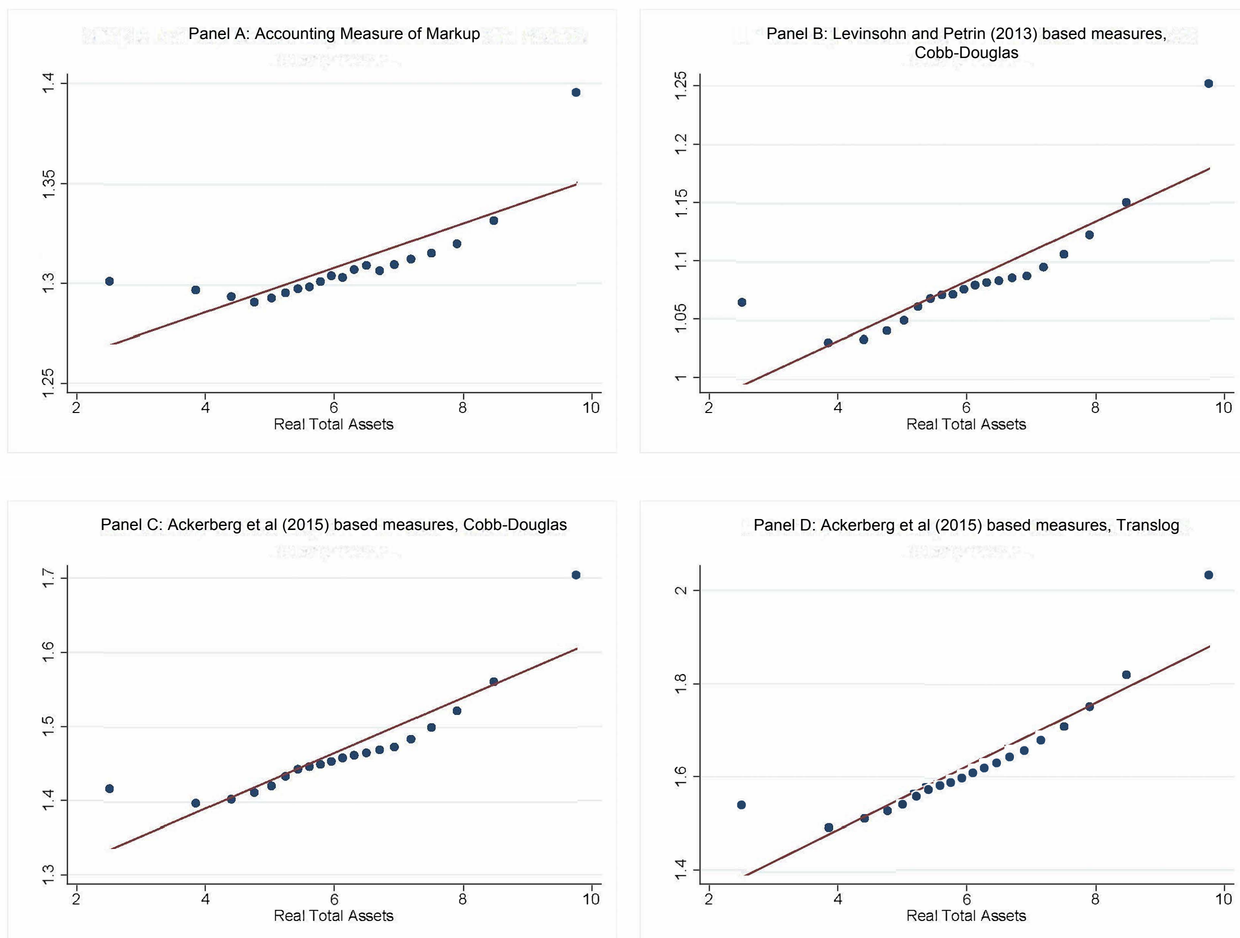
Notes. These are binscatters of firm TFP ($y-$axis) on the log of firm sales ($x-$axis), controlling for a full set of four-digit SIC industry by year dummies (so these are the relationships within an industry-year pair). Each panel has a different way of measuring TFP (the order of the panels follows that in Figure 10). The top left panel uses the accounting method which replaces output elasticities with the factor's share of total costs. The top right panel uses our plant-level, industry specific estimates of a Cobb-Douglas production function using the Levinsohn-Petrin method. The bottom left also uses estimates of a Cobb-Douglas production function, except using the Ackerberg et al (2015) method. The bottom right also uses Ackerberg et al (2015), but in the context of a translog production function. Labor services are measured by payroll, so each worker's input is weighted by their wage (as in Hsieh and Klenow, 2009).

Figure A.3: The Relationship between Labor Productivity and Firm Size by Sector



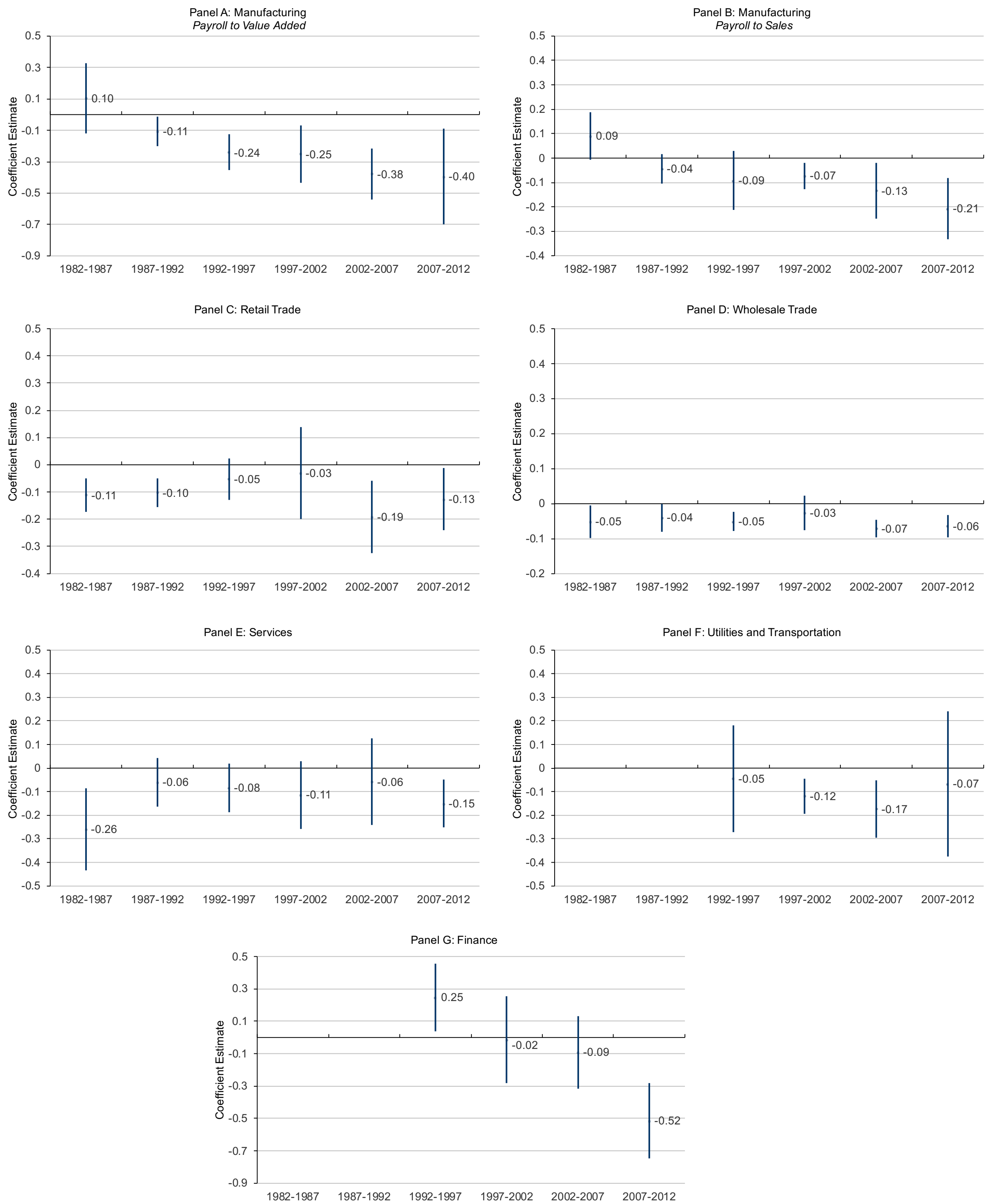
Notes. These are binscatters of firm ln(output per worker) on the $y-$axis and firm ln(sales) on the $x-$axis. We control for a full set of four-digit SIC industry by year dummies (so these are the relationships within an industry-year pair).

Figure A.4: The Relationship between Estimated Markups and Firm Size Using Four Methods of Estimating Markups



Notes. These are binscatters of firm markups on the $y-$axis and firm ln(capital) on the $x-$axis. We control for a full set of four-digit SIC industry by year dummies (so these are the relationships within an industry-year pair). The ordering of the panels follows Figure 10 in the main text.

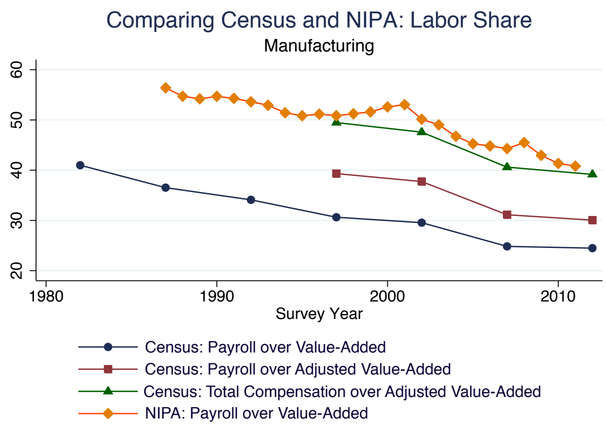
Figure A.5: Correlation Between the Change in Labor Share and the Change in Concentration: Period Specific Estimates



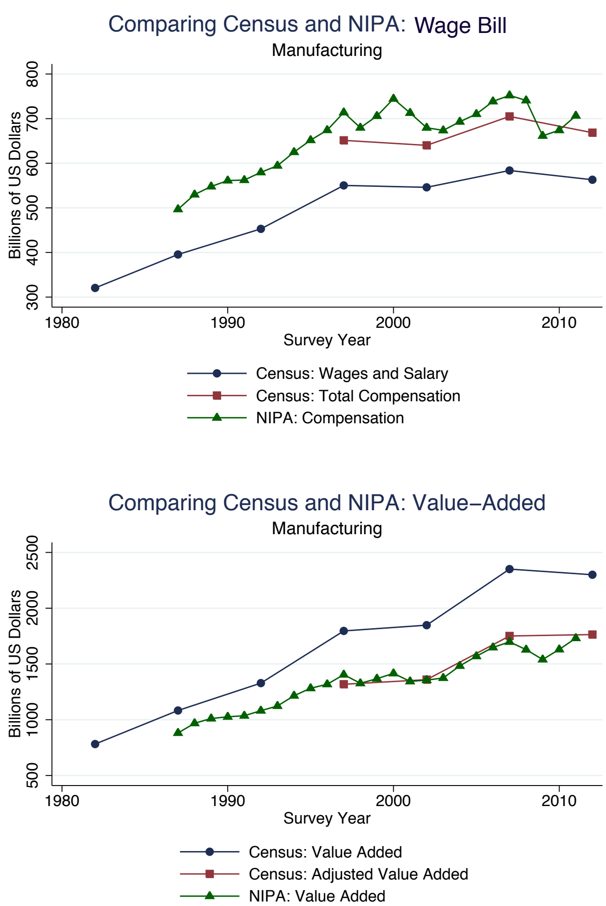
Notes. The labor share is defined using the payroll to value-added ratio in panel A, and each industry is weighted by the industry's 1982 share of value-added. For all other panels, the labor share is defined as the ratio of payroll to sales, and each industry is weighted by its initial share of sales in 1982 (except for the finance and utilities and transportation sectors, where initial sales shares are based on 1992 data due to shorter sample periods). Concentration is measured using CR20. Vertical lines represent the 95% confidence intervals.

Figure A.6: Comparing Labor Share in NIPA and Census: Manufacturing Only

Panel A: Labor Share

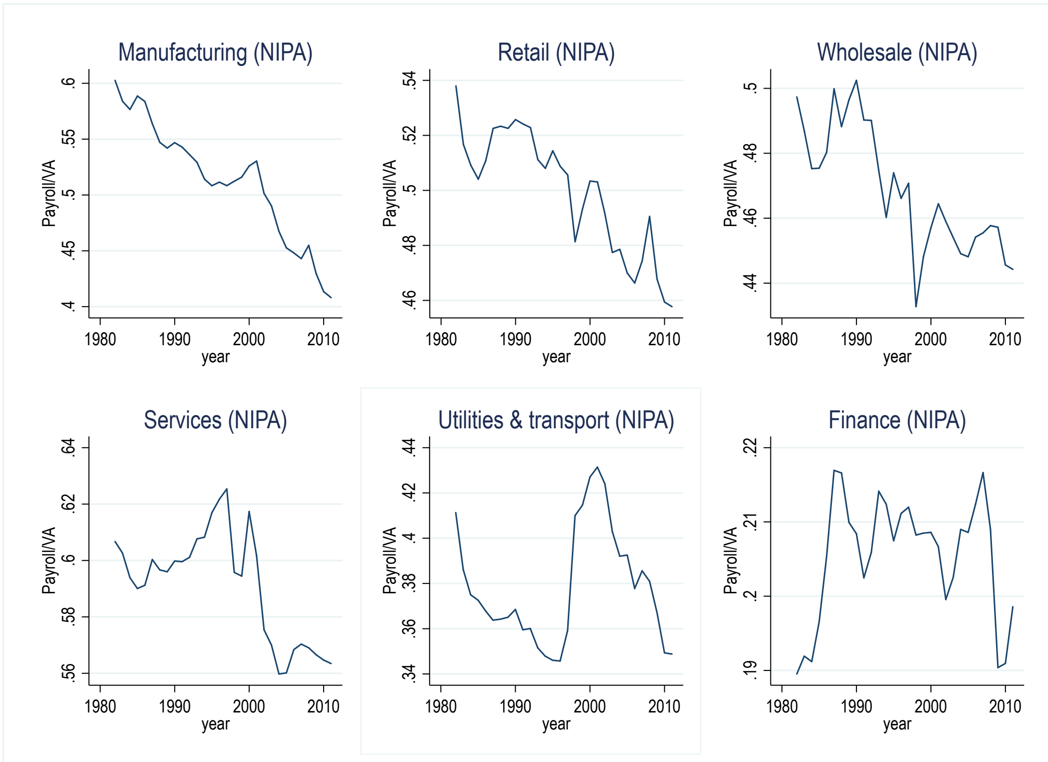


Panel B: Components of the Labor Share



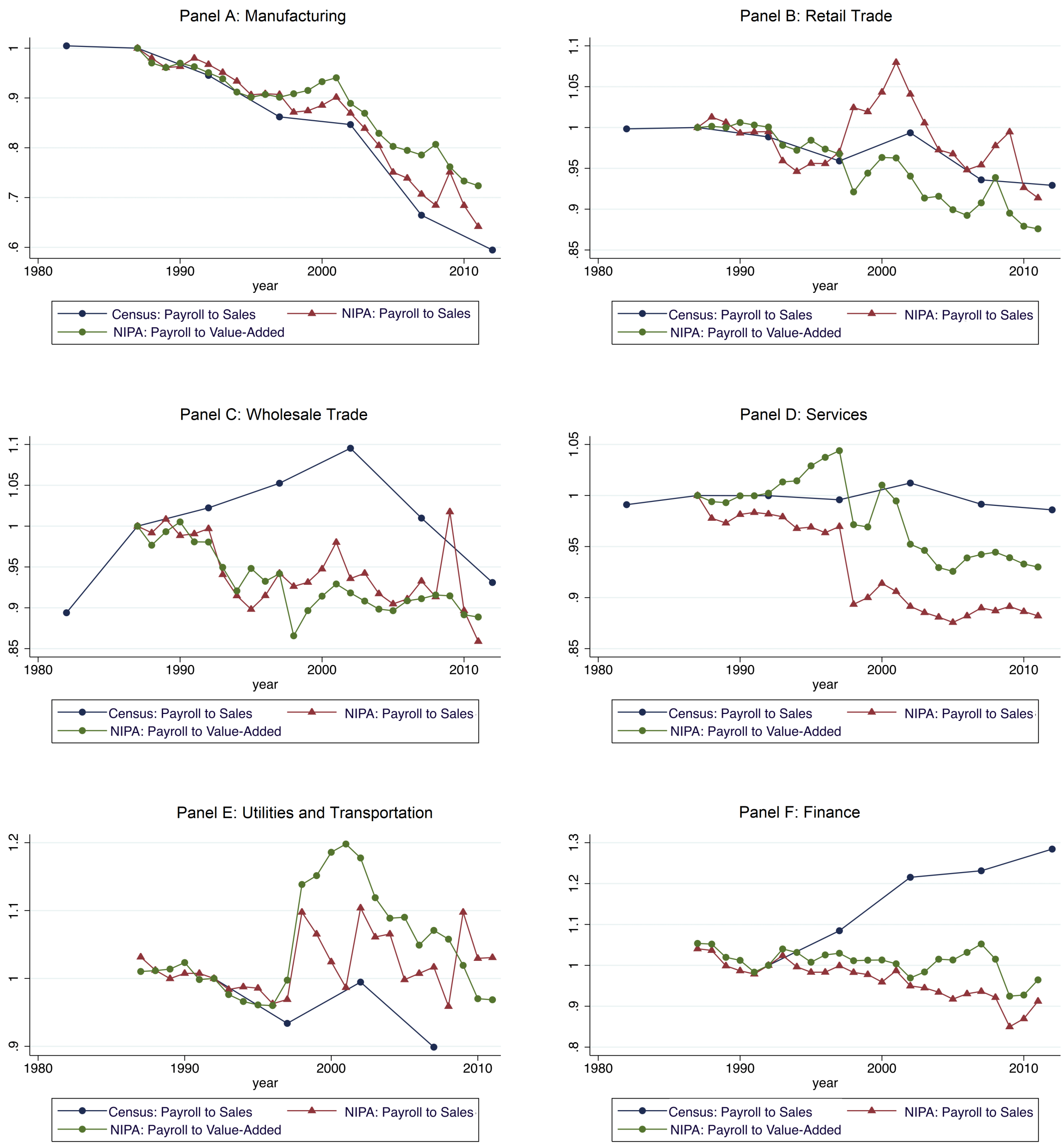
Notes. Panel A plots the aggregate labor share in Manufacturing calculated from the Census and NIPA/BEA data. Blue circles show the labor share calculated in the Census as the ratio of payroll to value-added. Red squares show the same ratio, but here value-added is adjusted by subtracting intermediate purchased services as described in Appendix B. Green triangles further augment the labor share to include additional labor costs to payroll. Lastly, the yellow diamonds plot the payroll over value-added from the NIPA data. Panel B separately plots the numerator (payroll) and denominator (value added) used in the construction of the labor shares in Panel A. These figures were from an initial disclosure and thus do not include the revised 2012 update of the Census.

Figure A.7: Labor Share in NIPA



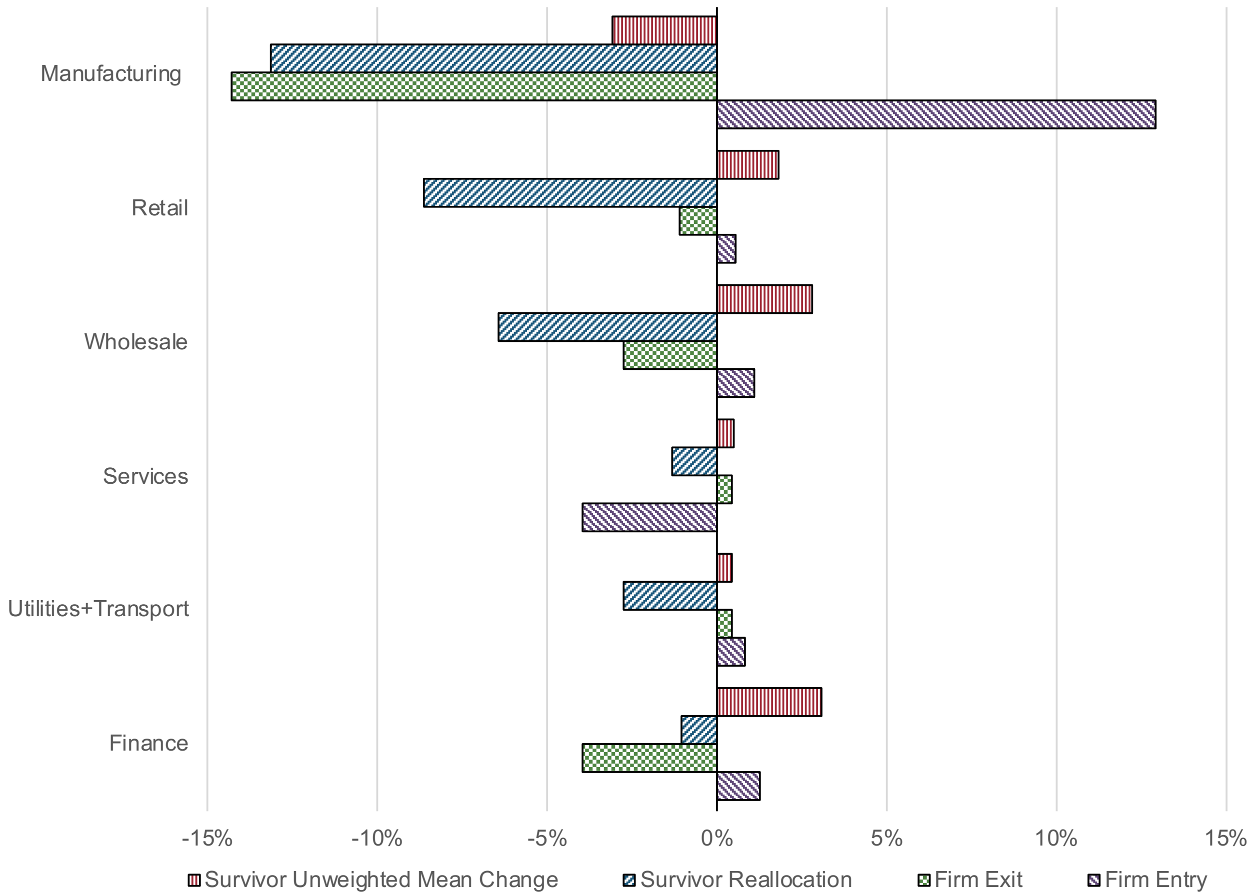
Notes. These are graphs of the ratio of payroll to value added taken from the NIPA/BEA data presented separately for each Census Sector. See text for details.

Figure A.8: Comparing the Payroll-to-Sales Ratio in the Census with the Labor Share in NIPA



Notes. Each panel shows the payroll to sales ratio in the Census, the payroll to gross-output ratio in the NIPA/BEA data, and the payroll to value-added ratio in the NIPA/BEA data. All series are normalized to one in 1987. These figures were from an initial disclosure and thus do not include the revised 2012 update of the Census.

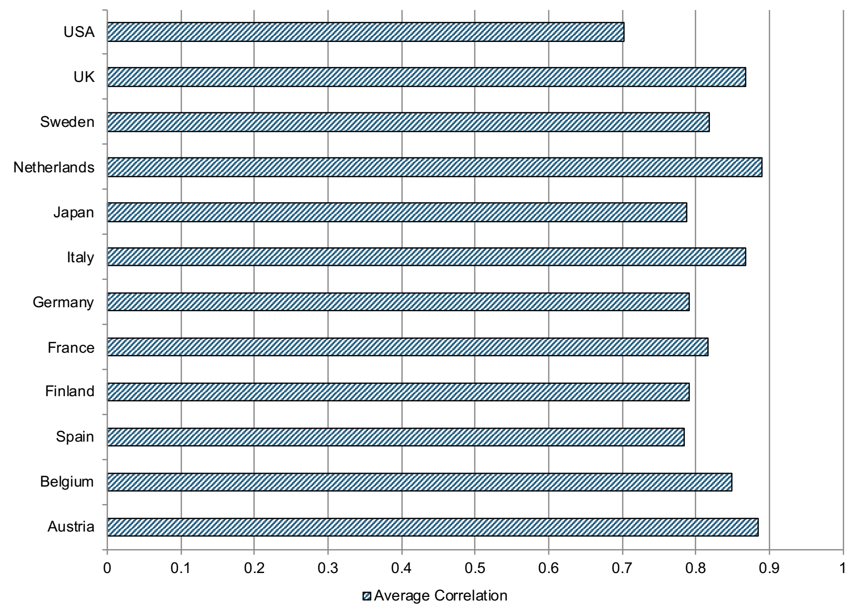
Figure A.9: Decomposition of the Labor Share Decline by Sector in the National Income and Product Accounts (NIPA)



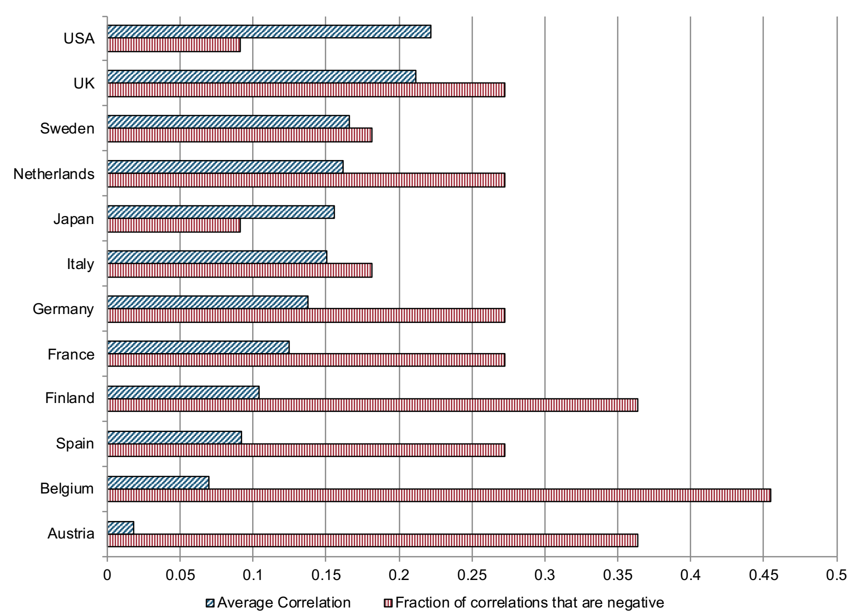
Notes. Melitz-Polanec decomposition of fall of labor share (payroll to value added) using NIPA and Census data. See text for details.

Figure A.10: Industry-Level Cross-Country Comparisons of Labor Shares

Panel A: Levels

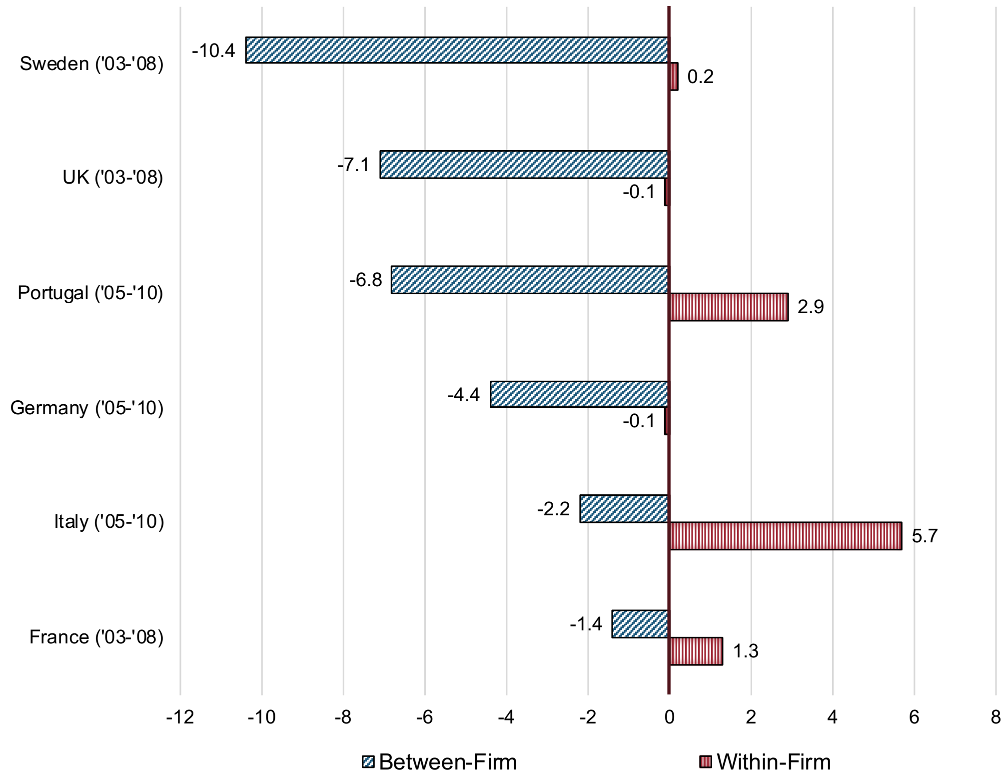


Panel B: Changes



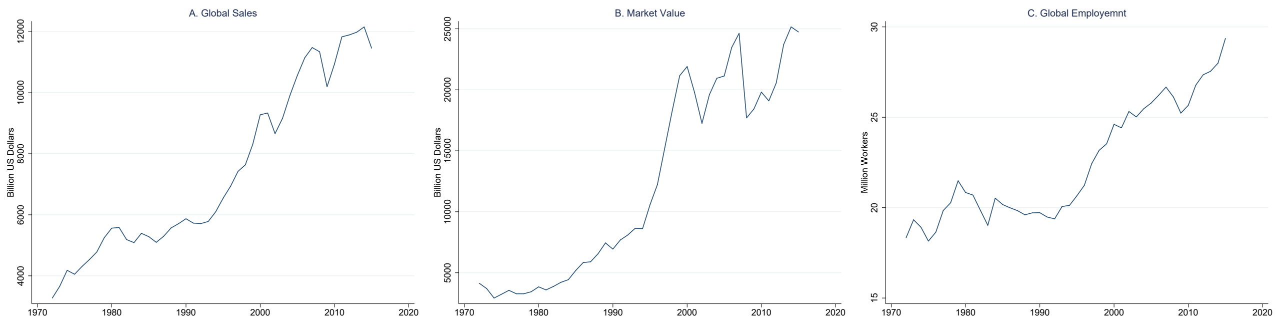
Notes. Using international KLEMS data, Panel A plots for each country the correlation of the *level* of its labor share in 32 industries with the corresponding industry-level labor shares in 11 other countries, averaged over the 11 pairwise correlations with each other country. Note that each cross-country correlation contributes twice to the calculation, as the correlation between the USA and the UK would enter the average correlation for the U.S. and the average correlation for the U.K. The light grey bars in Panel B plot the industry-level correlation of the ten-year *change* in the labor share, averaged over 11 country pairs. The darker solid bars in panel B show the fraction of the country pair correlations that are negative. The sample period in both panels is 1997-2007, and each industry in the correlation is weighted by the value-added share of that industry averaged over the two countries in comparison. In order to reduce measurement error, the correlations are calculated using centered five-year moving averages.

Figure A.11: Decomposing the Payroll Share Using Firm Level Data from Different Countries



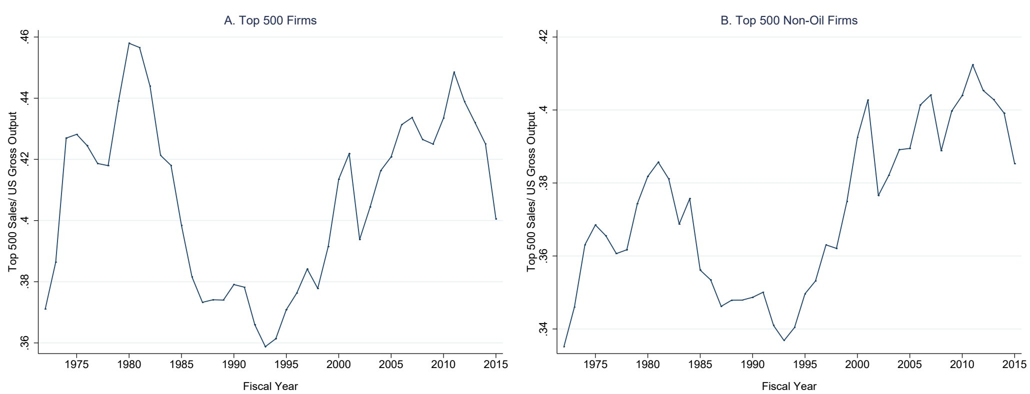
Notes. This figure plots Olley-Pakes decompositions of the change of the payroll share into between-firm and within-firm components (equation 4 in the text) using BVD Orbis Data. Between-firm refers to the reallocation component occurring between incumbent firms, while within-firm refers to the unweighted average change in the labor share. (BVD does not provide reliable data on entry and exit.) These calculations are performed over five-year periods within reliably-measured manufacturing data in indicated European countries. Labor share is payroll divided by value-added (equal to gross profits plus payroll). See Appendix for details of the firm-level panel data and exact numbers underlying the decompositions.

Figure A.12: Size of the Top 500 U.S. Firms



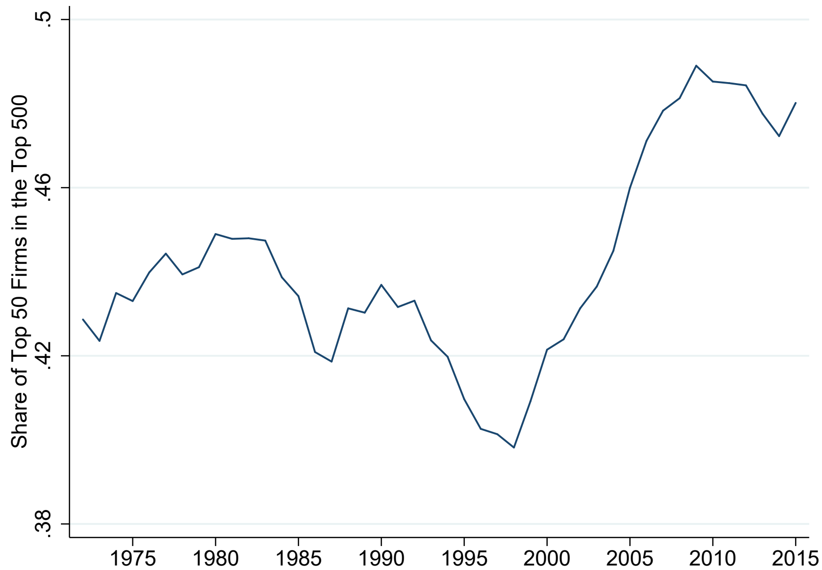
Notes. Panel A shows the total global sales for the 500 firms with the largest global sales from 1972 to 2015. Panel B shows the total market value for the 500 firms with the largest global sales from 1972 to 2015. Panel C shows the total global employment for the 500 firms with the largest global sales from 1972 to 2015. Sales and market value variables are inflated to 2015 using the GDP deflator.

Figure A.13: Ratio of Top 500 Firms' Sales to U.S. Gross Output



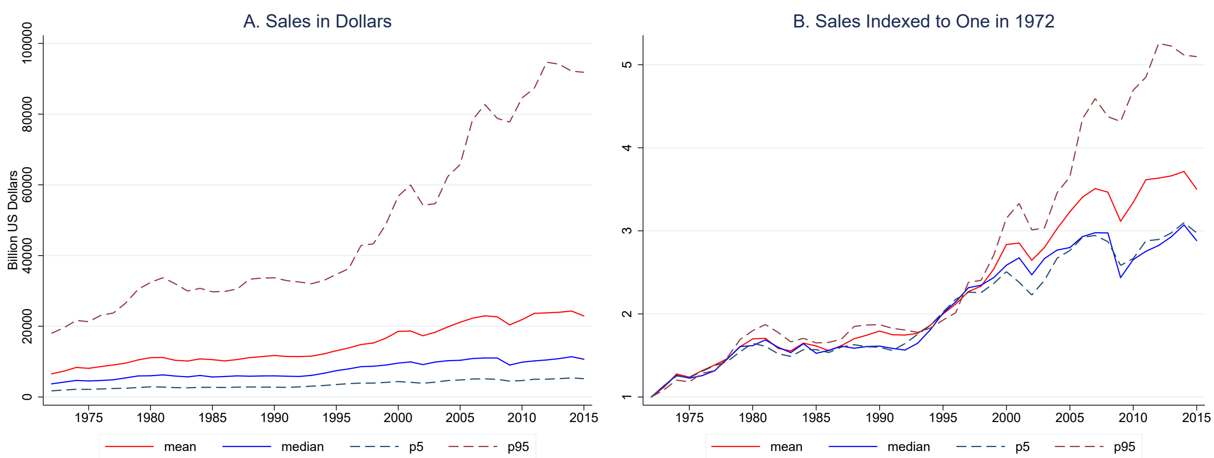
Notes. Panel A reports the ratio of aggregate global sales of the top 500 U.S. firms to total gross output of the U.S. private sector. Panel B reports the ratio of aggregate global sales of the 500 largest non-oil firms to total output of the non-oil private sector (omitting the oil and gas mining and petroleum refining industries).

Figure A.14: Share of Top 50 Firms in Combined Sales of Top 500 Firms



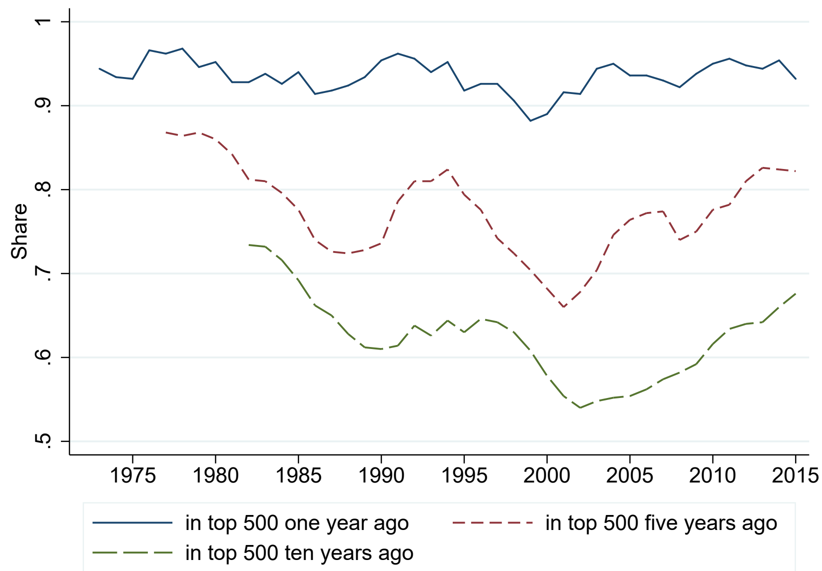
Notes. This numerator of this figure is the sum of global sales of the top 50 firms defined by global sales. The denominator is the sum of global sales of the top 500 firms defined by global sales.

Figure A.15: Quantiles of the Sales Distribution among the Top 500 Firms



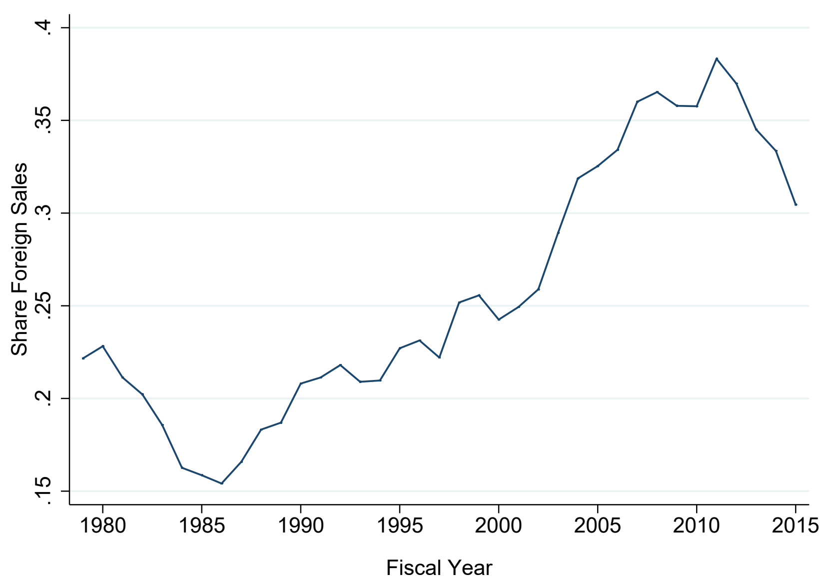
Notes. Panel A shows the time series from 1972 to 2015 for the mean, median, and 5th and 95th percentile of global sales among the top 500 largest firms defined by global sales. Panel B shows the same time series as in Panel A, with all series indexed to one in 1972. Dollar values are inflated to 2015 using the GDP deflator.

Figure A.16: Persistence of Firms in the Top 500



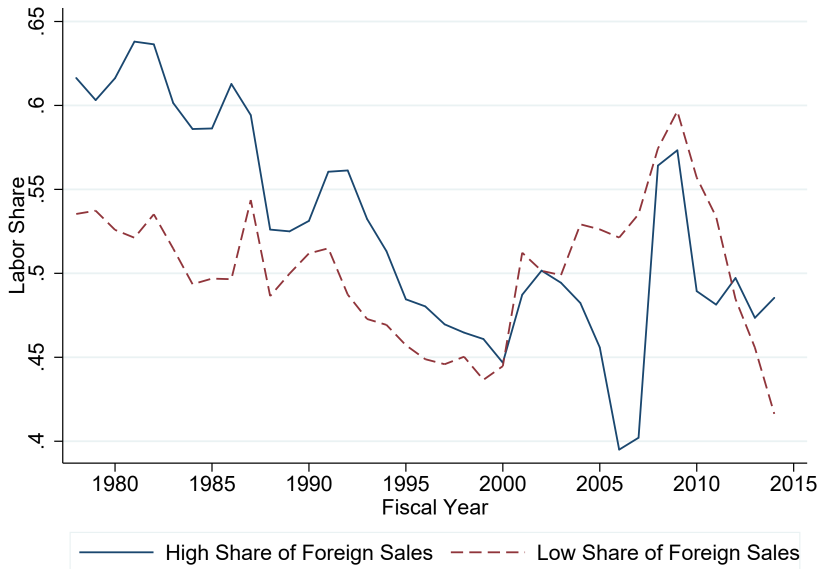
Notes. This figure plots the fraction of firms in the top 500 (defined by global sales) in the indicated year that were already in the top 500 1/5/10 years ago.

Figure A.17: Share of Foreign Sales among Top Firms



Notes. This figure shows the fraction of foreign sales in total sales among the top 500 firms defined by global sales.

Figure A.18: Labor Share among Top Firms by Extent of Global Engagement



Notes. This figure shows the average labor share separately for the top 500 firms by sales whose share of foreign sales in total sales is equal or larger than the median of their 2-digit industry (high share of foreign sales) and the firms with a foreign sales share below the industry median (low share of foreign sales). If there is only one top 500 firm in a given two-digit industry in a year, then the firm is classified as having a high share of foreign sales if its foreign sales share is equal or larger than the median for its broad sector (manufacturing or non-manufacturing). The average weights firms by sales, and omits firms for which the labor share cannot be measured in Compustat.

Appendix Tables

Table A.1: Decompositions of the Change in the Labor Share in Manufacturing: Alternative Aggregation Levels

	Wage Bill Share of Value Added				Compensation Share of Value Added			
	Δ Unweighted Mean	Incumbent Reallocation	Exit	Entry	Δ Unweighted Mean	Incumbent Reallocation	Exit	Entry
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. Plant Level				
1982-1987	-3.20	-0.85	-1.07	0.59	-3.20	-0.85	-1.07	0.59
1987-1992	2.32	-4.10	-1.12	0.31	2.32	-4.10	-1.12	0.31
1992-1997	-1.95	-1.42	-0.60	0.65	-1.95	-1.42	-0.60	0.65
1997-2002	0.51	-0.88	-0.75	0.03	0.51	-0.88	-0.75	0.03
2002-2007	-2.68	-1.58	-0.54	0.31	-2.68	-1.58	-0.54	0.31
2007-2012	2.34	-2.24	-0.36	0.17	2.34	-2.24	-0.36	0.17
1982-1997	-2.83	-6.37	-2.78	1.56	-2.83	-6.37	-2.78	1.56
1997-2012	0.18	-4.69	-1.65	0.51	0.18	-4.69	-1.65	0.51
1982-2012	-2.65	-11.06	-4.43	2.07	-2.65	-11.06	-4.43	2.07
				B. Firm by Industry Level				
1982-1987	-3.46	-1.07	-1.35	1.36				
1987-1992	2.19	-4.72	-1.57	1.52				
1992-1997	-2.24	-1.10	-1.54	1.57				
1997-2002	-0.06	-0.79	-1.47	1.25				
2002-2007	-3.00	-1.58	-1.82	1.92				
2007-2012	2.40	-2.29	-1.31	1.11				
1982-1997	-3.51	-6.89	-4.46	4.45				
1997-2012	-0.66	-4.67	-4.60	4.28				
1982-2012	-4.18	-11.56	-9.06	8.73				
				C. 15-Year Decompositions, Firm Level				
1982-1997	-3.79	-7.17	-1.58	2.18	-1.21	-12.07	-1.39	2.39
1997-2012	-2.29	-3.70	-2.08	1.91	-2.49	-4.03	-2.25	2.14

Notes. This table shows the results of a decomposition of the change in the labor share using the dynamic Melitz-Polanec methodology as described in the text. "Change in Unweighted Mean" is the change in the labor share due to a general fall in the share across all incumbent plants; "Incumbent Reallocation" is the change due to the growing relative size of low labor share incumbent plants; "Exit" is the contribution to the change from the exit of high labor share plants; and "Entry" is contribution from the entry of low labor share plants. All calculations use micro-data from the quinquennial Censuses of Manufacturing. Panel A reports the decomposition at the plant level, Panel B at the firm-by-industry level, and Panel C at the firm level over adjacent 15-year periods. An analysis of compensation share of value-added at firm level was not disclosed by the Census Bureau.

Table A.2: Decomposition of the Change in Payroll to Value-Added Ratio, Breaking Out Between- and Within-Industry Effects: Manufacturing Sector

		Industry Shift-Share		Within-Industry Melitz-Polanec Decomposition				
	Total	Between Industry Shifts	Within-Industry Changes	Δ Unweighted Mean	Incumbent Reallocation	Exit	Entry	Total Reallocation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. Payroll-to-Value added Ratio				
1982-87	-4.52	0.08	-4.60	-2.82	-1.56	-0.39	0.15	-1.79
1987-92	-2.58	-0.80	-1.78	1.16	-2.64	-0.60	0.30	-2.94
1992-97	-3.32	0.09	-3.41	-2.24	-1.08	-0.41	0.32	-1.17
1997-02	-1.08	-0.43	-0.65	1.07	-1.43	-0.49	0.21	-1.71
2002-07	-4.48	-0.70	-3.78	-3.29	-0.54	-0.53	0.57	-0.49
2007-12	-0.09	-0.48	0.39	1.06	-0.63	-0.33	0.30	-0.67
1982-12	-16.07	-2.23	-13.84	-5.07	-7.88	-2.76	1.87	-8.77
				B. Compensation-to-Value added Ratio				
1982-87	-5.93	0.03	-5.96	-2.41	-3.29	-0.44	0.18	-3.55
1987-92	-1.91	-0.96	-0.94	2.16	-2.62	-0.75	0.27	-3.10
1992-97	-4.64	0.16	-4.81	-2.80	-1.90	-0.44	0.33	-2.01
1997-02	-1.25	-0.55	-0.70	-0.24	-0.12	-0.60	0.26	-0.46
2002-07	-4.60	-1.01	-3.59	-0.69	-3.05	-0.58	0.73	-2.91
2007-12	-0.15	-0.54	0.39	-0.14	0.64	-0.40	0.29	0.53
1982-12	-18.48	-2.86	-15.62	-4.12	-10.35	-3.21	2.06	-11.50

Notes. This table reports an extended version of the decomposition as used in Table 4 where we break out the between four-digit SIC industry component from the within-industry component following equation (28) in Appendix D.

Table A.3: Decomposition of the Change in Payroll to Sales Ratio, Breaking Out Between- and Within-Industry Effects: All Sectors

	Ind Shift-Share		Within-Industry MP Decomp				Ind Shift-Share		Within-Industry MP Decomp			
	Between Industry Shifts	Within-Industry Changes	Δ Unweighted Mean	Incumbent Reallocation	Exit	Entry	Between Industry Shifts	Within-Industry Changes	Δ Unweighted Mean	Incumbent Reallocation	Exit	Entry
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
			A. Manufacturing						*B. Retail*			
1982-87	1.20	-1.21	-0.13	-1.02	-0.10	0.04	0.10	-0.10	-0.06	0.04	-0.02	-0.06
1987-92	-0.07	-0.98	0.96	-1.83	-0.18	0.07	0.06	-0.21	1.10	-1.12	-0.08	-0.11
1992-97	0.28	-1.61	-0.46	-1.15	-0.18	0.18	-0.17	-0.17	0.99	-1.11	-0.04	-0.01
1997-02	-0.19	-0.09	1.01	-0.99	-0.17	0.07	-0.01	0.37	1.36	-0.92	-0.05	-0.03
2002-07	-0.88	-2.11	-1.86	-0.22	-0.19	0.17	-0.06	-0.57	0.32	-0.79	-0.11	0.00
2007-12	-0.75	-0.33	-0.04	-0.26	-0.12	0.09	0.01	-0.25	0.76	-0.99	-0.06	0.03
1982-12	-0.41	-6.32	-0.52	-5.47	-0.95	0.61	-0.06	-0.94	4.47	-4.88	-0.35	-0.18
			C. Wholesale						*D. Services*			
1982-87	0.32	0.19	0.82	-0.61	-0.14	0.12	0.37	-0.34	-0.07	-0.08	0.66	-0.85
1987-92	0.22	-0.02	1.11	-0.94	-0.18	0.00	0.55	-1.06	0.23	-0.99	0.42	-0.72
1992-97	0.20	-0.01	1.32	-1.26	-0.12	0.05	0.27	-0.88	1.16	-2.07	0.41	-0.37
1997-02	-0.33	0.67	1.81	-1.28	-0.14	0.27	0.37	0.38	0.66	-0.32	0.28	-0.24
2002-07	-0.19	-0.33	0.05	-0.37	-0.16	0.15	0.20	-1.35	-0.58	-0.46	-0.21	-0.10
2007-12	-0.42	-0.29	0.91	-1.06	-0.25	0.11	0.25	-0.17	0.37	-0.28	-0.36	0.10
1982-12	-0.21	0.21	6.01	-5.52	-0.99	0.70	2.02	-3.40	1.77	-4.19	1.20	-2.18
			E. Utilities and Trnasportation						*F. Finance*			
1992-97	-0.24	-1.02	1.01	-1.96	0.08	-0.15	1.18	-0.17	2.00	-1.65	-0.28	-0.24
1997-02	0.88	0.26	0.57	-0.50	0.39	-0.21	0.17	1.21	0.92	0.36	-0.13	0.06
2002-07	-0.26	-1.62	-0.29	-1.47	0.65	-0.51	1.22	-1.15	1.02	-2.15	-0.03	0.01
2007-12	0.16	-0.32	0.30	-0.59	0.26	-0.29	-1.18	1.98	1.80	0.31	-0.13	0.00
1992-12	0.54	-2.71	1.58	-4.51	1.39	-1.17	1.40	1.88	5.74	-3.12	-0.57	-0.17

Notes. This table reports an extended version of the decomposition as used in Table 4 where we break out the between four-digit SIC industry component from the within-industry component following equation (28) in Appendix D.

Table A.4: **Output Elasticities for Production Function Estimates**

		LP				ACF			
		K		L		K		L	
Industry	Obs	(1)		(2)		(3)		(4)	
Food and kindred products	85500	0.260	***	0.472	***	0.345	***	0.693	***
Textile mill products	21000	0.164	***	0.611	***	0.183	***	0.796	***
Apparel and other textile products	81500	0.213	***	0.523	***	0.293	***	0.627	***
Lumber and wood products	86000	0.191	***	0.600	***	0.210	***	0.803	***
Furniture and fixtures	44000	0.168	***	0.558	***	0.293	***	0.833	***
Paper and allied products	31500	0.213	***	0.557	***	0.229	***	0.796	***
Printing and publishing	134000	0.171	***	0.631	***	0.209	***	0.823	***
Chemicals and allied products	55000	0.253	***	0.428	***	0.330	***	0.663	***
Petroleum and coal products	12500	0.223	***	0.383	***	0.341	***	0.664	***
Rubber and misc. plastics products	66000	0.196	***	0.543	***	0.230	***	0.750	***
Leather and leather products	6600	0.173	***	0.525	***	0.201	***	0.785	***
Stone, clay, and glass products	75500	0.228	***	0.468	***	0.247	***	0.709	***
Primary metal industries	31500	0.180	***	0.620	***	0.219	***	0.804	***
Fabricated metal products	163000	0.160	***	0.648	***	0.193	***	0.817	***
Industrial machinery and equipment	194000	0.144	***	0.670	***	0.219	***	0.863	***
Electronic & other electric equipment	54000	0.165	***	0.575	***	0.193	***	0.818	***
Transportation equipment	34500	0.175	***	0.607	***	0.204	***	0.836	***
Instruments and related products	44500	0.197	***	0.566	***	0.214	***	0.813	***
Miscellaneous manufacturing industries	58500	0.167	***	0.564	***	0.217	***	0.793	***

Notes. The table reports output elasticities from industry specific estimates of the production function. The analysis uses plant-level panel data from the Census of Manufactures 1982-2012. Columns (1) and (2) apply the Levinsohn and Petrin (2003) method, while columns (3) and (4) use the Ackerberg, Caves and Frazer (2015) method. Both are based on Cobb-Douglas approaches with time trends. See text for further details. * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A.5: Top 25 Largest Publicly Listed U.S. Firms by Global Sales in 1985, 2000 and 2015

	1985			2000			2015		
Rank	Company	Industry	Sales $bn	Company	Industry	Sales $bn	Company	Industry	Sales $bn
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	General Motors Co	Automobiles	185.5	Exxon Mobil Corp	Petroleum	277.3	Wal-Mart Stores Inc	Merch. Stores	476.6
2	Exxon Corp	Petroleum	166.9	Wal-Mart Stores Inc	Merch. Stores	252.8	Exxon Mobil Corp	Petroleum	237.6
3	AT&T Corp	Telecom	108.6	General Motors Co	Automobiles	242.9	Apple Inc	Computers	234.5
4	Mobil Corp	Petroleum	107.7	Ford Motor Co	Automobiles	228.8	Berkshire Hathaway	Conglomerate	211.5
5	Ford Motor Co	Automobiles	101.6	General Electric Co	Conglomerate	172.3	McKesson Corp	Drugs Wholes.	189.5
6	IBM Corp	Computers	96.4	Citigroup Inc	Banking	150.5	Unitedhealth Group	Insurance	157.6
7	Texaco Inc	Petroleum	89.1	Enron Corp	Energy	135.6	CVS Health Corp	Pharmacies	153.8
8	Chevron Corp	Petroleum	80.4	IBM Corp	Computers	118.9	General Motors Co	Automotive	152.8
9	Sears Roebuck &Co	Dept Stores	78.4	AT&T Corp	Telecom	88.8	Ford Motor Co	Automotive	150.0
10	Du Pont Co	Chemicals	56.4	Verizon Comm. Inc	Telecom	87.2	AT&T Inc	Telecom	147.3
11	General Electric Co	Conglomerate	54.5	Altria Group Inc	Tobacco	85.1	AmerisourceBergen	Drugs Wholes.	136.4
12	Travelers Group	Insurance	53.7	JPMorgan Chase	Banking	79.3	Verizon Comm. Inc	Telecom	132.0
13	Amoco Corp	Petroleum	51.8	Bank of America	Banking	77.9	Chevron Corp	Petroleum	123.0
14	Citicorp	Banking	43.3	SBC Comm. Inc	Telecom	69.3	Costco Wholesale	Merch. Stores	116.6
15	Kmart Corp	Merch. Stores	42.6	Boeing Co	Airplanes	69.1	General Electric Co	Conglomerate	115.5
16	Atlantic Richfield	Petroleum	41.8	Texaco Inc	Petroleum	67.4	The Kroger Co	Food Stores	109.1
17	Chrysler Corp	Automobiles	40.9	Duke Energy Corp	Oil and Gas	65.8	Amazon.com Inc	Internet Sales	107.3
18	Shell Oil Co	Oil and Gas	39.1	HP Inc	Computers	65.6	Walgreens Boots	Pharmacies	103.8
19	Safeway Inc	Food Stores	37.8	The Kroger Co	Food Stores	64.5	HP Inc	Computers	103.7
20	Aetna Inc.	Insurance	35.8	Chevron Corp	Petroleum	62.6	Cardinal Health Inc	Drugs Wholes.	102.9
21	USX Corp	Steel	35.5	AIG Inc	Insurance	61.9	Express Scripts Co	Pharma Serv.	102.1
22	The Kroger Co	Food Stores	33.0	Morgan Stanley	Banking	61.1	JPMorgan Chase	Banking	100.8
23	Cigna Corp	Insurance	31.2	Merrill Lynch & Co	Banking	60.4	Boeing Co	Aircraft	96.4
24	GTE Corp	Telecom	30.3	Home Depot Inc	Hardware St.	60.2	Microsoft Corp	Software	93.9
25	Phillips Petroleum	Petroleum	30.1	Compaq Computer	Computers	57.0	Bank of America Co	Banking	93.4
	Total Sales Top 25 Firms		1,672	Total Sales Top 25 Firms		2,762	Total Sales Top 25 Firms		3,748

Notes. Dollar values are inflated to 2015 using the GDP deflator.

Table A.6: Regressions of the Components of the Change in the Payroll-to-Sales Ratio on the Change in Concentration

	CR4 (1)		CR20 (2)		HHI (3)	
			A. Incumbent Reallocation			
Retail	-0.038	**	-0.071	***	-0.038	
Wholesale	-0.013		-0.023	*	-0.038	
Services	-0.167	***	-0.186	***	-0.434	***
Manufacturing	-0.063	***	-0.087	***	-0.08	**
Utilities/Transportation	-0.102	*	-0.122	**	-0.325	***
Finance	-0.247		-0.237	**	-0.543	**
Combined	-0.077		-0.086	***	-0.119	***
			B. Change in Unweighted Mean			
Retail	0.003		0.005		0.005	
Wholesale	-0.016	*	-0.005		-0.016	
Services	0.066	***	0.078	***	0.104	**
Manufacturing	0.015	*	0.037	***	-0.003	
Utilities/Transportation	-0.014		-0.020		-0.014	
Finance	0.001		-0.034		-0.038	
Combined	0.005		0.006		-0.010	
			C. Entry			
Retail	0.007		-0.015		0.021	
Wholesale	-0.008	**	-0.010	***	-0.017	
Services	0.001		0.010		-0.007	
Manufacturing	-0.004		-0.019	***	-0.004	
Utilities/Transportation	0.034	***	0.038	***	0.077	**
Finance	0.034	**	0.044	*	0.045	
Combined	0.004		-0.002		0.006	
			D. Exit			
Retail	-0.007		-0.003		-0.029	**
Wholesale	-0.001		-0.001		-0.012	
Services	0.017		-0.022		-0.001	
Finance	-0.010		-0.008		-0.025	
Manufacturing	-0.028	*	-0.007		-0.056	*
Utilities/Transportation	-0.008		-0.025	**	-0.031	*
Combined	-0.008	*	-0.006	*	-0.027	**

Notes. Numbers of observations as indicated for models with five-year changes in Table 3. Each cell displays the coefficient from a separate OLS industry-level regression of the change in labor share (payroll-to-sales ratio) on period fixed effects and a component of the decomposition of changes in concentration as in Table (6). Industries are weighted by their sales in the initial year, and standard errors in parentheses are clustered by four-digit industries. * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A.7: Industry-Level Cross-Country Comparisons of Labor Shares

	Outcome Mean	Austria	Belgium	Spain	Finland	France	Germany	Italy	Japan	Nether-lands	Sweden	UK	USA
						A. Correlations of Average Value-Added Share Levels, 1997 - 2007							
Austria	65.47	1.00											
Belgium	64.95	0.73	1.00										
Spain	67.17	0.54	0.93	1.00									
Finland	65.79	0.73	0.81	0.75	1.00								
France	64.90	0.82	0.93	0.85	0.91	1.00							
Germany	64.52	0.75	0.87	0.71	0.80	0.89	1.00						
Italy	63.82	0.75	0.94	0.87	0.84	0.95	0.89	1.00					
Japan	62.89	0.63	0.81	0.80	0.81	0.82	0.75	0.82	1.00				
Netherlands	65.33	0.76	0.82	0.75	0.81	0.92	0.67	0.88	0.78	1.00			
Sweden	62.93	0.45	0.87	0.82	0.82	0.82	0.77	0.81	0.83	0.79	1.00		
UK	65.91	0.74	0.92	0.83	0.83	0.93	0.76	0.89	0.80	0.90	0.80	1.00	
USA	62.48	0.82	0.93	0.87	0.88	0.95	0.80	0.90	0.86	0.92	0.86	0.95	1.00
						B. Correlations of 10-Year Value-Added Share Levels, 1997 - 2007							
Austria	-0.60	1.00											
Belgium	-0.61	-0.06	1.00										
Spain	-0.48	0.22	0.45	1.00									
Finland	-0.53	0.53	0.22	0.48	1.00								
France	-0.36	-0.23	0.43	0.40	0.11	1.00							
Germany	-0.88	0.52	-0.15	-0.20	0.23	-0.13	1.00						
Italy	-0.53	0.09	0.28	0.26	-0.16	0.13	-0.02	1.00					
Japan	-1.02	0.23	0.29	-0.09	0.15	0.04	0.27	0.08	1.00				
Netherlands	-1.00	0.11	0.24	0.01	0.23	0.29	0.24	0.09	0.23	1.00			
Sweden	-0.93	0.16	-0.02	-0.39	0.28	0.06	0.28	-0.34	0.29	0.11	1.00		
UK	-1.06	-0.14	0.44	0.53	0.06	0.51	-0.39	0.08	-0.08	-0.10	0.06	1.00	
USA	-1.02	0.07	0.21	0.11	0.32	0.22	0.11	-0.31	-0.05	0.27	0.53	0.18	1.00

Notes. Correlations include 32 industries both within and outside of manufacturing, using international KLEMS data. In each correlation, industries are weighted by their value-added share over the two countries in the comparison. Panel A correlates labor share levels, averaged between 1997 and 2007. Panel B correlates 10-year changes in labor share between 1997-2007. To reduce measurement error, we estimate correlations using centered five-year moving averages.

Table A.8: International COMPNET Regressions of the Change in Labor Share on the Change in Concentration (Industry level, all sectors)

	5 Year Δ		10 Year Δ		Obs
	(1)		(2)		(3)
Italy	-0.124	**	-0.200	**	53
	(0.052)		(0.095)		
Estonia	-0.140		-0.125		53
	(0.197)		(0.084)		
Portugal	-0.083		---		53
	(0.063)		---		
Slovenia	-0.106		-0.101		53
	(0.140)		(0.187)		
Slovakia	-0.153	**	-0.343	***	52
	(0.060)		(0.100)		
Finland	-0.208	***	-0.181	**	53
	(0.059)		(0.076)		
Belgium	-0.008		0.330	*	53
	(0.053)		(0.176)		
Germany	-0.091		-0.151		44
	(0.060)		(0.094)		
Poland	0.007		---		53
	(0.076)		---		
France	0.325		-0.183	**	53
	(0.255)		(0.087)		
Latvia	-0.039		---		52
	(0.108)		---		
Romania	-0.137		---		53
	(0.096)		---		
Austria	-0.297	***	-0.275	**	37
	(0.098)		(0.108)		
Lithuania	-0.124		-0.045		53
	(0.156)		(0.201)		

Notes. Concentration is defined at the fraction of industry output produced by the ten largest firms. Regression includes five-year changes for 2006-2011 and ten-year changes (when available) for 2001-2011. Observations are weighted by the industry's share of the country's total value-added. Models are estimated by OLS with standard errors clustered at the industry level.

Table A.9: Decomposing the Wage Bill Share Using Firm-Level Data from Different Countries

	Period	Obs	Initial Labor Share	Δ Labor Share	Δ Unweighted Mean	Incumbent Reallocation	Exit	Entry
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
UK	2003-08	118,584	69	-7.6	-0.1	-7.1	-2.6	2.2
Sweden	2003-08	174,252	74	-2.8	0.2	-10.4	7.2	0.3
France	2003-08	752,400	76	-1.7	1.3	-1.4	-1.6	-0.1
Germany	2005-10	140,567	76	-4.6	-0.1	-4.4	-0.3	0.2
Italy	2005-10	929,287	73	2.6	5.7	-2.2	-1.0	0.1
Portugal	2005-10	248,905	71	-4.8	2.9	-6.8	-1.9	1.0

Notes: This uses firm data from BVD Orbis. Value added is constructed by adding wage bill to pre-tax profits (EBIT). Firms whose primary three digit industry is in manufacturing. We use the MP method to break down the aggregate change into a between and within firm component.

Notes. The table uses firm-level data from BVD Orbis. Value-added is constructed by adding wage bill to pre-tax profits (EBIT) for firms whose primary three-digit industry is in manufacturing. We use the MP method to break down the aggregate change into a between- and within-firm component.

Table A.10: The Labor Share and the Rise in Chinese Imports

Δ Years	ln(Sales) (1)	ln(Payroll) (2)	ln(Value-Added) (3)	CR4 (4)	CR20 (5)	HHI (6)	Labor Share (7)	Payroll-to-Sales (8)
				A. OLS Estimates				
5 Year Δ's	-1.98 **	-0.46 *	-0.79 **	1.16	0.341	1.18	6.64 **	2.28
1992-2012	(0.77)	(0.28)	(0.35)	(4.39)	(4.12)	(2.00)	(2.98)	(1.82)
10 Year Δ's	-2.55 ***	-0.83 **	-1.36 ***	-4.89	-1.80	-0.85	12.38 ***	6.85 ***
1992-2012	(0.76)	(0.34)	(0.43)	(7.91)	(7.30)	(3.64)	(2.98)	(1.14)
5 Year Δ's	-2.66 ***	-0.66 **	-0.67 **	16.58 *	7.36 **	11.17 **	-1.44	-1.10
1992-2007	(1.00)	(0.30)	(0.26)	(9.23)	(3.25)	(5.39)	(2.98)	(1.12)
				B. 2SLS Estimates				
5 Year Δ's	-3.72 ***	-0.78 **	-1.17 ***	4.69	3.50	4.80	8.17 **	3.60 **
1992-2012	(1.41)	(0.34)	(0.42)	(5.24)	(4.01)	(3.17)	(3.30)	(1.79)
10 Year Δ's	-4.10 ***	-1.21 ***	-1.93 ***	-3.15	3.47	2.03	15.77 ***	8.42 ***
1992-2012	(1.26)	(0.43)	(0.56)	(9.34)	(7.13)	(5.38)	(3.30)	(1.61)
5 Year Δ's	-2.66 ***	-1.05 ***	-1.17 ***	17.60 *	9.84 **	13.12 **	0.52	-0.95
1992-2007	(1.00)	(0.38)	(0.40)	(9.57)	(4.49)	(6.20)	(3.30)	(1.42)

Notes. N=1,552 in rows 1 and 4, N=776 in rows 2 and 5, and N=1,164 in rows 3 and 6 (388 manufacturing industries x 4/2/3 periods). Each cell displays the coefficient from a separate regression of the change in the variable indicated at the top of the column on the change in Chinese import penetration and period fixed effects. Industries are weighted by their total value added in 1982, and standard errors in parentheses are clustered by industries. Panel A reports OLS regressions while Panel B reports 2SLS estimates using the growth in imports from China to eight other developed countries as an instrument for the contemporaneous growth in Chinese imports to the U.S. (as in Autor et al. 2019). * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.