A Appendix

A.1 Additional Stylized Facts

The relationships discussed in Section 3 are robust to a broad set of controls such as plant size and age, industry effects, unionization of production workers, foreign-ownership, military production, using value added instead of revenues, or whether or not a plant exports. Tables 3(a)-3(b) show the estimated coefficients of the technology index conditional on these controls.[27] The results indicate that, controlling for observables, a 1 percent increase in technology index is associated with a 0.04-0.08 percent decline in production labor share, 0.12-0.14 percent increase in production labor productivity, and 0.08-0.09 percent increase in average production worker wage across plants. In contrast, the technology index does not seem to be related to non-production labor share, whereas average wage and labor productivity of non-production workers both increase with the technology index.[28] Overall, these results confirm the bivariate relationships discussed above and are robust when value-added-based measures of labor share and labor productivity are used. Unreported results indicate correlation between some plant characteristics and production labor share. For example, younger, larger, domestically-owned, and intensely-exporting plants generally have lower production labor share. In addition, younger and larger plants rely more heavily on automation.

A.2 Evolution of SMT Industries

In order to shed more light on the general evolution of some key indicators for the SMT industries, Figures A3(a)-A3(c) show industry-level factor share measures for the five industries over the period 1972-1997.[29] Both overall and production labor shares decline in the 5 industries during this period, see Figures A3(a)-A3(b). The capital share is relatively stable across all industries, with some increase observed in SIC 34 and 36, see Figure A3(c). These dynamics are reflected in the ratio of capital share to production labor share, see Figure A3(d). The 5-factor TFP measures in Figure A3(e) suggest that productivity is higher and grows faster in more capital intensive industries (notably, SIC 35 and 36). Overall, the trends are broadly similar and in line with the general decline in labor share in manufacturing.

A.3 Optimization

Given σ and $(\alpha_i|\sigma)$, K_i and L_{pi} are determined. The interior solution to cost minimization satisfies the following first-order conditions: $w_{ji}X_{ij}{=}\lambda^*\beta_jQ_i$, $w_{ki}K_i{=}\lambda^*Q_i\gamma\alpha^{\frac{2}{\sigma}}K_i^{\rho}T_i^{\gamma-1}$, $w_{pi}L_{pi}{=}\lambda^*Q_i\gamma(1-$

[27] For robustness, these tables also show Technology Index II that is based on the average investment indicator across the four technology groups based on survey question 2 in Table 1(b).

[28] Non-production worker category includes labor with various education and skill levels. The confounding effect of this unobserved heterogeneity may explain the greater standard errors for its share.

[29] We use industry-level data from the NBER-CES Manufacturing Industry Database, available at http://www.nber.org/nberces/, supplemented with capital prices from the Bureau of Labor Statistics. The descriptive statistics about the inputs and output in these industries can be found in Table A5.

$\alpha)^{\frac{2}{\sigma}} L_{pi}^{\rho} T_i^{\gamma-1}$, $K_i^{\rho} \alpha^{\frac{2}{\sigma}-1}$=$(1-\alpha)^{\frac{2}{\sigma}-1} L_{pi}^{\rho}$, where λ^* denotes the Lagrange multiplier and w_{ji} denote factor prices. These conditions imply that the cost function can be written as TC_i=$\sum_j w_{ji} X_{ji} = \lambda^* Q_i \left(\sum_j \beta_j + \gamma\right)$. The second order condition for cost minimization with respect to α_i holds when $\sigma < 1$. The first order conditions can be solved for the capital labor ratio and the relative weight in equations (3)-(4). Similarly, the first order conditions of profit maximization for the jth variable input can be written as w_{ji}=$\beta_j \frac{Q_i}{X_{ji}}$. For K_i and L_{pi} these read $\frac{Q_i}{K_i} \gamma \alpha_i^{2/\sigma} K_i^{\rho} T_i^{\gamma-1}$=$w_{ki}$, and $\frac{Q_i}{L_{pi}} \gamma (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho} T_i^{\gamma-1}$=$w_{pi}$.

A.4 Details of Production Function Estimation

1. Use (11) to obtain an estimate of σ.

2. Estimate variable input elasticities using $\widehat{\beta}_j$=$N^{-1} \sum_i \frac{w_{ji} X_{ji}}{R_i}$.

3. Subtract variable input costs from revenues: $\widehat{\mathcal{B}}_i$=$\ln R_i - \sum_j \widehat{\beta}_j \ln (w_{ji} X_{ji})$.

4. Conditional on $\widehat{\alpha}_i = \frac{w_{ki} K_i}{w_{ki} K_i + w_{pi} L_{pi}}$ and $\widehat{\rho} = \frac{\widehat{\sigma}-1}{\widehat{\sigma}}$, calculate the log-composite input as

$$\ln \widehat{T}_i \equiv \frac{1}{\widehat{\rho}} \ln \left[\widehat{\alpha}_i^{2/\widehat{\sigma}} K_i^{\widehat{\rho}} + (1-\widehat{\alpha}_i)^{2/\widehat{\sigma}} L_{pi}^{\widehat{\rho}}\right]. \tag{14}$$

5. Determine the joint contribution of state variables and the proxy by estimating

$$\widehat{\mathcal{B}}_i = \phi(Z_i, p) + v_i, \tag{15}$$

   where $\phi(Z_i, p)$ denotes a polynomial of degree p in vector Z_i, which contains state variables and the proxy. Choosing $p = 2$ is standard. The vector of state variables includes $\ln \widehat{T}_i$ and other plant characteristics, such as plant age. If the only state variable is $\ln \widehat{T}_i$ and if automation investment can be subsumed into a scalar indicator $\bar{t}_i$, then $Z_i = \left(\ln \widehat{T}_i, \bar{t}_i\right)'$.[30]

6. Given fitted values $\widehat{\phi}_{it}$ from equation (15), use nonlinear least squares to estimate

$$\widehat{\mathcal{B}}_{it} = \delta_T \ln \widehat{T}_{it} + h\left(\widehat{\phi}_{it-1} - \delta_T \ln \widehat{T}_{it-1}\right) + \nu_{it}. \tag{16}$$

   where h is a second-order polynomial in its argument. Assuming a Markovian data generating process for productivity, as in Olley and Pakes (1996), $h\left(\widehat{\phi}_{it-1} - \delta_T\right)$ approximates $E[\ln \theta_{it} | \ln \theta_{it-1}]$, and δ_T in regression (16) identifies γ.

As mentioned earlier, the SMT asks about the plant's total investment in technologically advanced equipment and software for the previous three years for each of the four technology groups. The proxy, $\bar{t}_i$, is calculated as the average response of plants over the four technology groups. Together

[30] Treating T_i as a state variable can be justified by the considerations that lead to treating K_i as a state variable in the vast majority of the empirical productivity literature. Differences in establishments' productivity histories are controlled for by the proxy $\bar{t}_i$, if the only unobservable is productivity and if investment in technology is an increasing function of productivity.

with $\ln \widehat{T}_i$, and plant age they determine $\widehat{\phi}_i$ in (15). Under the assumptions of the model, this value controls for unobserved productivity differences across plants when estimating δ_T using data from 1992 in (16). If the plant-level productivity process is Markovian –a standard assumption in the empirical productivity literature– then δ_T is consistently estimated in regression (16). The standard error of $\widehat{\delta}_T$ is obtained using the bootstrap.

A.5 Specification Error

This appendix studies the properties of the specification error when TFP is based on a standard Cobb-Douglas (CD) production function with constant returns to scale

$$Q_i = \theta_i^{\text{CD}} K_i^{\beta} L_{pi}^{1-\beta}, \tag{17}$$

when the underlying data generating process is a CES production function with endogenous technology choice

$$Q_i = \theta_i^{\text{CES}} T_i^{\gamma} \tag{18}$$

where T_i=$[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}]^{1/\rho}$ and γ<1. For ease of exposition, both production functions abstract from variable inputs shown in equation (2). Including them does not change the main conclusions in this section.[31]

Let k_i=lnK_i, l_{pi}=lnL_{pi}, $\widehat{\sigma}$=1/(1-$\widehat{\rho}$), $\text{CD}_{\text{CRS},i}$=ln$\widehat{\theta}_i^{\text{CD}}$ and $\text{CES}_{\text{EN},i}$=ln$\widehat{\theta}_i^{\text{CES}}$, and denote the log-difference between the two productivity residuals by Δ_i. In this notation Δ_i is given by

$$\Delta_i = \text{CD}_{\text{CRS},i} - \text{CES}_{\text{EN},i} = \frac{\widehat{\gamma}}{\widehat{\rho}} \ln[\alpha_i^{2/\widehat{\sigma}} K_i^{\widehat{\rho}} + (1-\alpha_i)^{2/\widehat{\sigma}} L_{pi}^{\widehat{\rho}}] - \widehat{\beta} k_i - (1-\widehat{\beta}) l_{pi},$$

which shows the log-difference between the two productivity residuals can be written as the difference between the two input indices, see equation (12) in the main text. Δ_i can be parsed into three components:

$$\begin{aligned}
\Delta_i = {} & \frac{\gamma}{\rho} \ln[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}] - \beta k - (1-\beta) l_{pi} \\
& + \frac{\widehat{\gamma}}{\widehat{\rho}} \ln[\alpha_i^{2/\widehat{\sigma}} K_i^{\widehat{\rho}} + (1-\alpha_i)^{2/\widehat{\sigma}} L_{pi}^{\widehat{\rho}}] - \frac{\gamma}{\rho} \ln[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}] \\
& + (\widehat{\beta} - \beta)(k_i - l_{pi}),
\end{aligned}$$

where the expression in first line can be thought of as the contribution by functional form assumptions. We denote it by Δ_i^S, where subscript S is short for "specification". The expressions in the second and third lines can be thought of as contributions by the estimation error associated with the CES and CD specifications. We denote them as $\Delta_{i,CES}^E$ and $\Delta_{i,CD}^E$, respectively. In this

[31] However, including variable inputs can matter if their elasticities are estimated using different methods for the CES and CD specifications.

notation, Δ_i can be written as

$$\Delta_i = \Delta_i^S + \Delta_{i,CES}^E + \Delta_{i,CD}^E \tag{$\star$}$$

Equation ($\star$) implies that Δ_i^S closely approximates Δ_i if for some positive number ε

$$\left|\Delta_{i,CES}^E\right|, \left|\Delta_{i,CD}^E\right| < \varepsilon \tag{$\star\star$}$$

holds. Equation ($\star\star$) can be thought of as a condition for the estimation error to be small, which seems reasonable given the results in Table 5.[32]

Assuming condition ($\star\star$) holds, we can rewrite equation ($\star$) as

$$\Delta_i \approx \Delta_i^S = \frac{\gamma}{\rho} \ln[\alpha_i^{2/\sigma} K_i^\rho + (1-\alpha_i)^{2/\sigma} L_p^\rho] - \beta k_i - (1-\beta) l_{pi} \tag{19}$$

and study Δ_i^S by evaluating he right hand side of equation (19) at different values of K/L and parameters. Before doing that, we develop some intuition behind the findings and the end of Section 6.2 by considering the CD and CES input indices

$$\text{CD} = K^\beta L^{1-\beta}$$
$$\text{CES} = [\alpha K^\rho + (1-\alpha) L^\rho]^{\gamma/\rho}$$

where σ=1/(1-ρ).[33] For simplicity, let α,β=1/2, γ=1, and 0<σ<1, which imply CD is equivalent to the geometric mean of inputs and CES is equivalent to the harmonic mean of inputs. This terminology is useful because then it follows from the properties of Pythagorean means that, all else equal, CES<CD and consequently Δ_i^S<0. In words, a Cobb-Douglas input index is greater than what technology-related complementarities between K and L would imply. Therefore, holding output constant, CES implies higher productivity residuals relative to CD.

We explore the properties of Δ_i^S by evaluating the right hand side of equation (19) as a function of K, L and α, conditional on (β,γ,σ). The K/L-dimension is defined as K, $L \in [0.9, 300]$, i.e. a 100-point equidistant grid, which is mapped in the into the interval $[-5.81, 5.81]$ on logarithmic scale. The α-dimension is defined as a 100-point equidistant grid in the interval $[0.001, 0.999]$. We choose (β,γ,σ)=$(1/2, 0.22, 0.6)$ as the reference parameter vector. β=1/2 provides an example for an industry with equal industry cost-shares for K and L, γ=0.22 is our baseline γ-estimate, σ=0.6 is a value consistent with the range of estimates from the SMT and ASM. We study the behavior of Δ_i^S by comparing baseline results to those calculated under β=0.1, β=0.9, γ=0.44, σ=0.6. The small (large) β-value is intended to illustrate Δ_i^S in a hypothetical industry with small (large) industry capital cost share. The additional γ and σ values are useful for studying Δ_i^S's sensitivity with respect to key parameters of the CES input index.

Figure A4(a) shows Δ_i^S under the baseline parameter vector. Δ_i^S is: negative for all K/L and

[32]Note that $\Delta_i \approx \Delta_i^S$ under a milder condition: $\left|\Delta_{i,CES}^E - \Delta_{i,CD}^E\right| < \varepsilon$, which holds under equation ($\star\star$).
[33]Note that $\lim_{\sigma,\gamma\to 1}$ CES=CD.

α; the smallest where $|\alpha$-$\beta|$ is the largest; the greatest for extreme K/L values. Figure A4(b) shows that a smaller β amplifies these effects for small K/L values and reverses Δ_i^S's sign for large K/L values. Figures A4(c)-A4(d) show that γ increases the curvature of Δ_i^S and more so for high α. Figures A4(e)-A4(f) illustrate that σ increases curvature by shifting up Δ_i^S around α=1/2 and this effect is virtually symmetric with respect to this contour line. The right hand side panels of Figure A5 show that a greater β reverses the relationship between Δ_i^S and K/L.

We develop intuition using these visualizations. Given that α can be calculated as the share of capital costs in the CES composite cost and it is correlated with Technology index I, see equation (6) and Figure 3(b), we can say that a business is an above-average (below-average) automator if its K/L is high (low), and its α is high (low). The point corresponding to this business is located in the quadrant of (K/L, α)-space that is closest to (furthest from) the viewer's perspective in Figures A4-A5. If an above-average automator operates in an industry with low industry capital cost share (β), Figure A4(b) is relevant. In this case, Δ_i^S is smallest for high K/L and becomes more negative as α_i increases. For a below-average automator in the same industry, Δ_i^S is closer to zero and eventually becomes positive as α increases. This is intuitive because the production process of the first business is better described by CES while that of the second one is closer to CD. Conversely, if an above-average automator operates in an industry with high β, Figure A5(b) is relevant. In this case, Δ_i^S is close to zero but it becomes eventually negative as α_i increases. For a below-average automator in the same industry, Δ_i^S is negative and remains negative as α grows. It is useful to note the similar shape of but opposite slopes on these two figures. Specifically, the derivative of the hyperplane with respect to of K/L is generally positive in Figures A5(b)- A5(f), and the derivative with respect to α is negative in subsets of the K/L-α space. These properties will be important when interpreting results in Table A1.

We note in the next-to-last paragraph of Section 6.2 that Δ_i<0 for the majority of establishments in the SMT. This inference is drawn from inspecting the empirical distribution of this variable (not shown here). It is useful to compare the properties of Δ_i^S with results from regressing Δ_i on other outcomes, shown in Table A1. Such a comparison is helpful because these partial correlations are based on the joint distribution of α_i, Δ_i, other plant-level outcomes and estimated parameters. Columns 1-2 indicate that Δ_i is decreasing in both Technology index I and establishment size, and column 3 shows Δ_i is increasing in K/L in our sample. Column 4 shows these relationships hold in a multivariate regression, which – together with the fact that K/L is greater than 1 and increasing in Technology index I in the SMT (see Figure 3(c) in the main text) – suggest that the general shape of Δ_i in the SMT could have sections similar to the 1<K/L-sections of the hyperplanes in Figures A4(a) and A5(b) because these have subsets that decrease in α and increase in K/L for 1<K/L.

A.6 Additional Results

Table A1: The relationship between Δ_i and plant characteristics

	$\Delta = \text{CD}_{\text{CRS}} - \text{CES}_{\text{EN}}$					
	I	II	III	IV	V	VI
Technology index I	-0.321***			-0.036***		
	[0.012]			[0.007]		
Employment		-0.250***		-0.255***	-0.251***	-0.245***
		[0.003]		[0.003]	[0.003]	[0.003]
Capital/production labor			0.064***	0.113***		
			[0.008]	[0.004]		
Technology index I					0.056***	0.055***
× capital/produciton labor					[0.012]	[0.012]
R^2	0.16	0.69	0.02	0.75	0.69	0.70
N	4600	4600	4600	4600	4600	4600

Notes: All regressions are estimated by OLS, see notes to Table 3. Specification VI includes other plant characteristics aside from employment.

Table A2: The relationship between the change in production labor share and automation with survival bias correction

	Growth in Production Labor Share			
	1997	2002	1997	2002
Technology index I	-0.080***	-0.085***		
	[0.015]	[0.019]		
Technology index II			-0.078***	-0.077***
			[0.015]	[0.020]
Employment growth (1997)	0.133***		0.130***	
	[0.012]		[0.013]	
Employment growth (2002)		0.170***		0.169***
		[0.012]		[0.012]
Mills λ	-0.005	-0.078*	-0.008	-0.074*
N	8100	8100	8100	8100

Notes: The coefficient estimates are based on the Heckman two-step correction. All continuous variables in logs and inverse hyperbolic sine transformation is used for the technology index. Standard errors are clustered by 4-digit SIC industry. (*), (**), (***) indicate significance at 10, 5, and 1% level, respectively. Technology index I is based on all 4 survey questions in Table 1(b). Technology index II is based only on the investment question (Question 2). Second-step includes the plant characteristics listed in Table 3(a). First-step includes, in addition, a dummy variable for whether the plant belongs to a multi-unit firm. N is rounded for disclosure avoidance.

Table A3: The relationship between the change in production labor productivity and automation with survival bias correction

	Growth in Production Labor Productivity			
	1997	2002	1997	2002
Technology index I	0.096***	0.090***		
	[0.014]	[0.019]		
Technology index II			0.105***	0.086***
			[0.014]	[0.020]
Employment growth (1997)	-0.168***		-0.168***	
	[0.012]		[0.012]	
Employment growth (2002)		-0.157***		-0.156***
		[0.012]		[0.012]
Mills λ	-0.015	0.137***	-0.004	0.136***
N	8100	8100	8100	8100

Notes: See notes to Table A2.

Table A4: The relationship between production labor share, automation and productivity

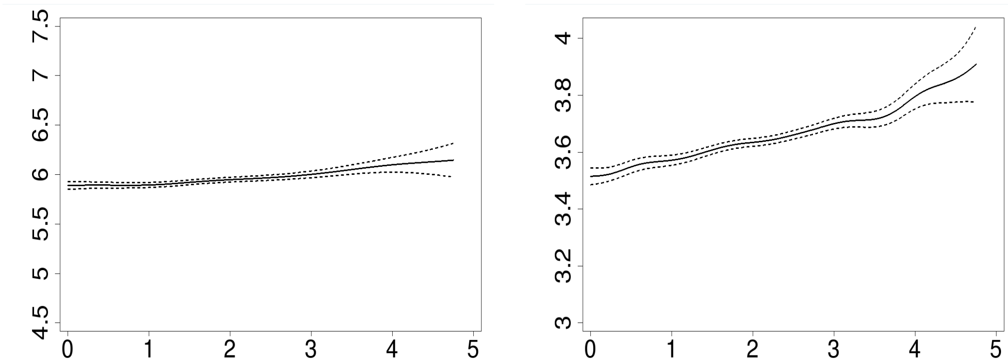
	Production labor's share in		Technology index	
	Revenue	Composite input expenditures	I	II
CES_{EN}	-0.177***	0.184***	0.279***	0.360***
	[0.028]	[0.013]	[0.013]	[0.013]
R^2	0.02	0.07	0.10	0.17
CD_{CRS}	-0.648***	0.012	0.076***	0.117***
	[0.034]	[0.020]	[0.022]	[0.021]
R^2	0.11	0.0001	0.003	0.008
N	4600	4600	4600	4600

Notes: The coefficient estimates are based on bivariate regressions. All continuous variables in logs and inverse hyperbolic sine transformation is used for the technology index. Standard errors are clustered by 4-digit SIC industry. (*), (**), (***) indicate significance at 10, 5, and 1% level, respectively. CES_{EN} and CES_{CRS} denote productivity residuals described in Section 6.2. Technology index I is based on all 4 survey questions in Table 1(b). Technology index II is based only on the investment question (Question 2). All variables are expressed as deviations from 4-digit SIC industry means. The productivity measures are averages over 1991 and 1992 by plant. For each dependent variable, the corresponding cells include the estimated coefficient of the productivity measure, its standard error and R^2, in that order. N is rounded for disclosure avoidance.

Table A5: Descriptive statistics of SMT industries, 1991

Industry	SIC Code	Prod. Lab. Sh.	Non-Prod. Lab. Sh.	Cap. Sh.	Cap. Share/ Prod. Lab. Sh.	Avg. TFP (5-factor)
Fabricated Metal	34	0.150	0.084	0.062	0.415	0.938
Industrial Machinery	35	0.116	0.109	0.059	0.507	0.993
Electronic & Other Electric	36	0.101	0.109	0.125	1.243	0.984
Transportation	37	0.095	0.070	0.025	0.263	0.996
Instruments & Related	38	0.092	0.158	0.038	0.415	1.032
Average, SMT industries		0.111	0.106	0.062	0.568	0.989
Average, Non-SMT industries		0.106	0.058	0.057	0.745	0.988

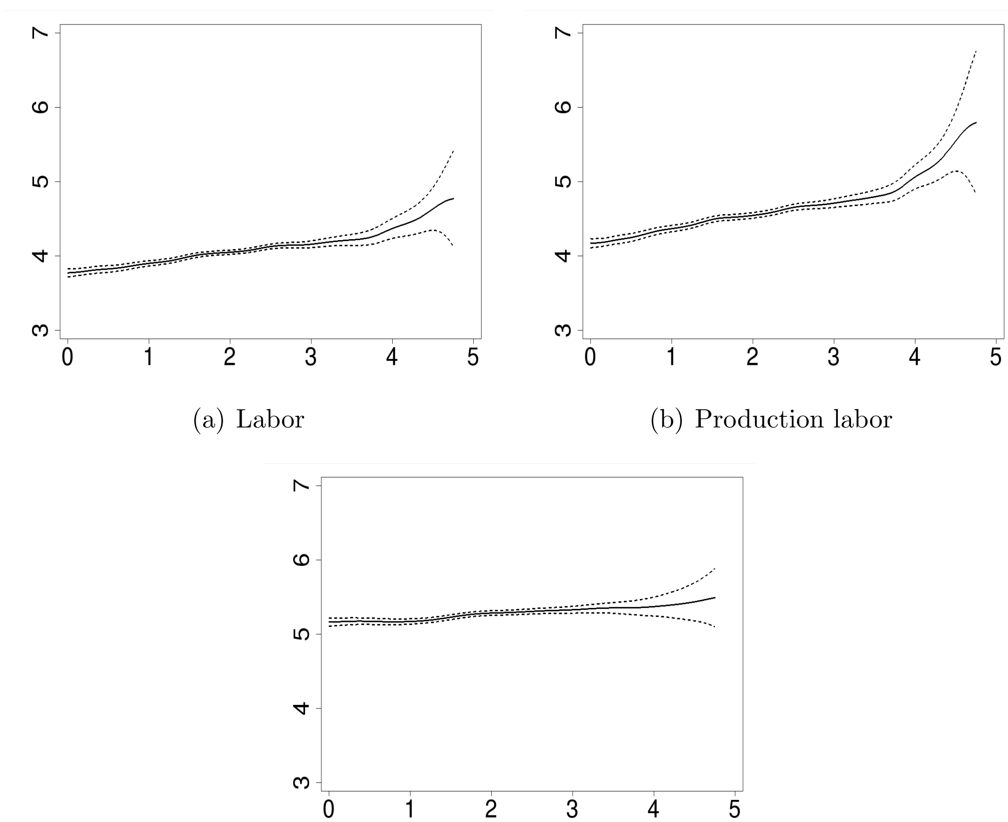
Source: NBER-CES Manufacturing Productivity Database. Average TFP is calculated across 4-digit industries within each 2-digit industry. Non-SMT average is based on 2-digit SIC industries outside the five SMT industries.



(a) Labor Productivity (b) Average Wage

Figure A1: Non-production labor productivity, average wage (in logs) and technology.

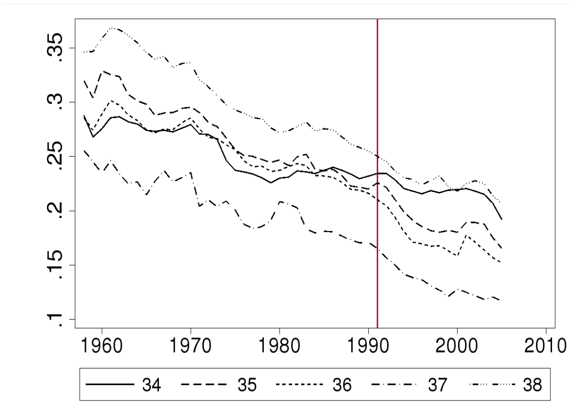
Notes: Non-parametric local polynomial estimates. The horizontal axes show "Technology Index I" described in Section 2.1. Variables on vertical axes are calculated using data from the Census Bureau's Collaborative Micro-Productivity Project, see section 2.2. Dotted lines show the 95% confidence intervals of smoothed local polynomial estimates.



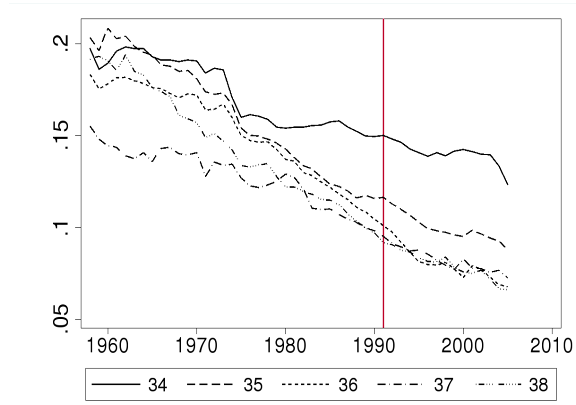
(a) Labor (b) Production labor

(c) Non-production labor

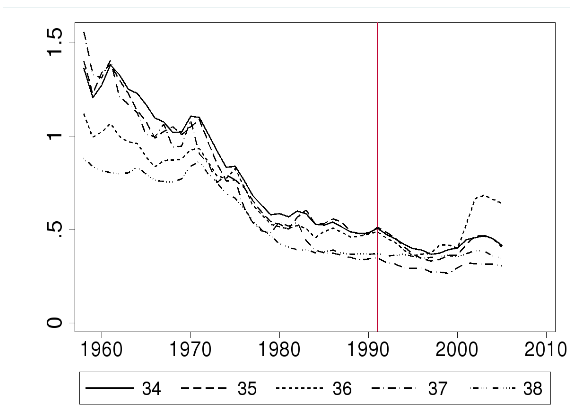
Figure A2: Value added per worker (in logs) and technology, see notes to Figure A1.



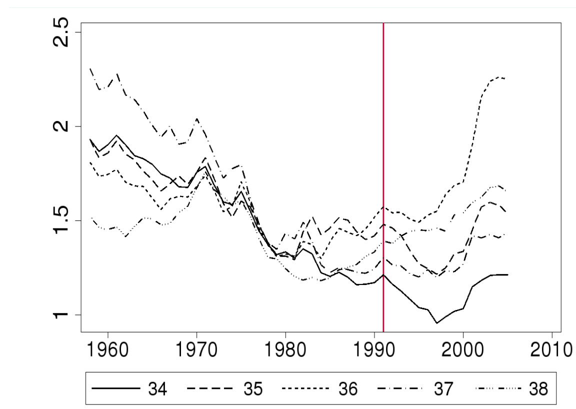
(a) Labor



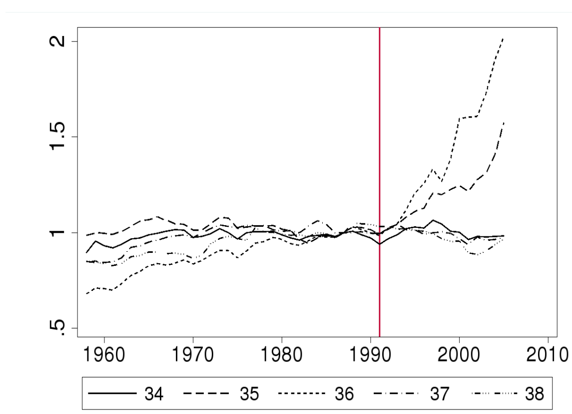
(b) Production labor



(c) Capital



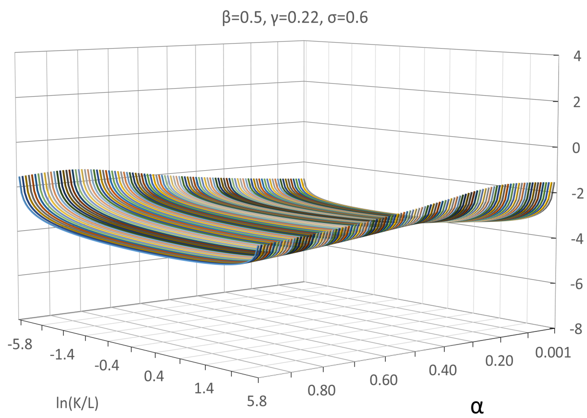
(d) Ratio of capital cost share to production labor cost share



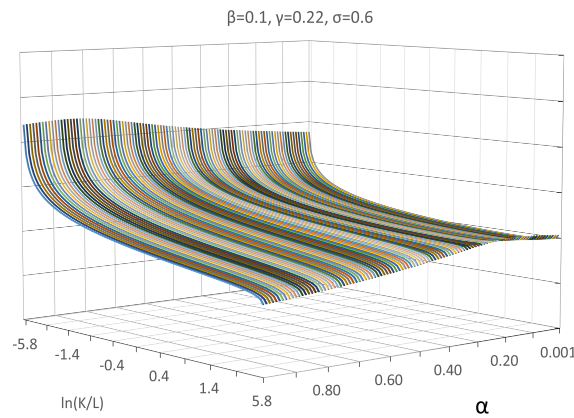
(e) Average log-TFP

Figure A3: The shares of capital and labor costs in the total value of shipments and industry TFP.

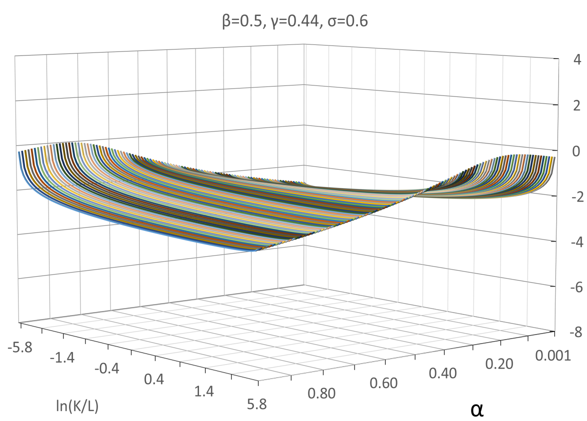
Source: NBER-CES database. Two-digit SIC codes denote industries included in the Survey of Manufacturing Technology: Fabricated Metal Products (34), Industrial Machinery and Equipment (35), Electronic and Other Electric Equipment (36), Transportation Equipment (37), Instruments and Related Products (38). Vertical lines mark SMT survey year 1991. The capital stock measure is not quality-adjusted.



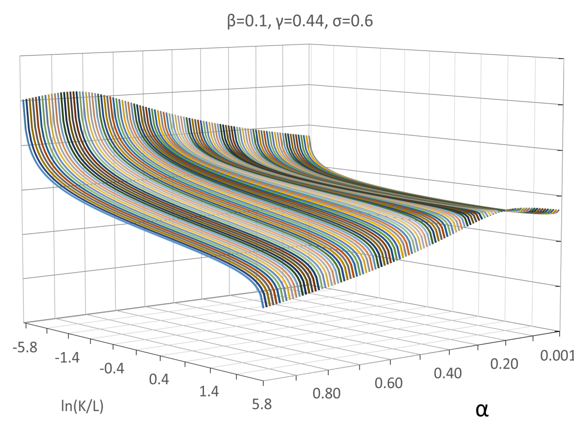
(a) Industry with equal CD elasticities



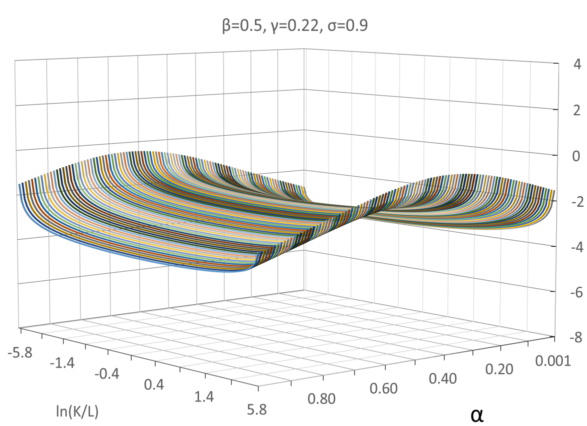
(b) Small β_k



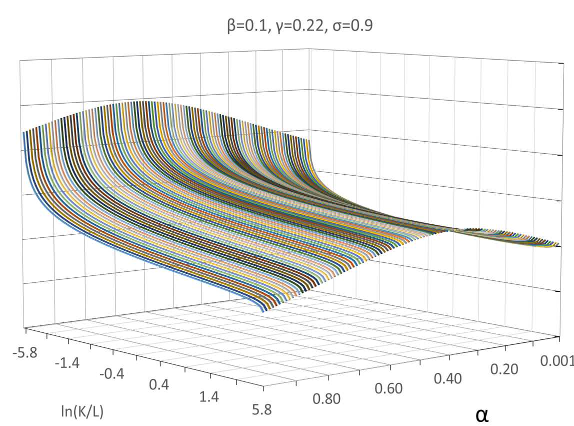
(c) Greater γ



(d) Small β_k, greater γ



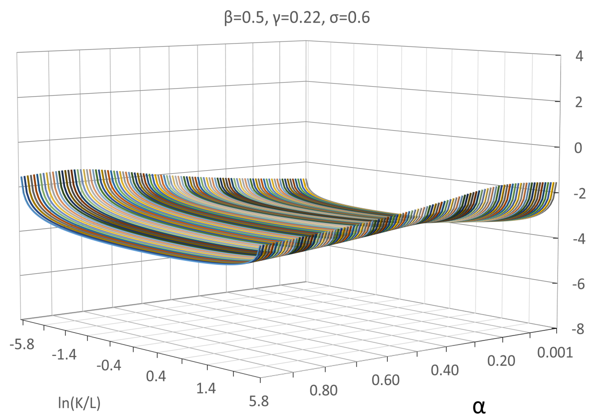
(e) Greater σ



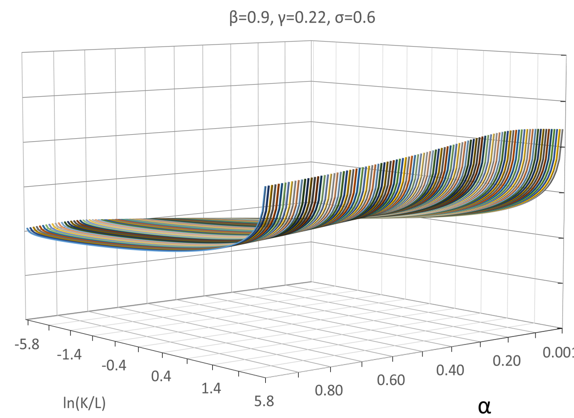
(f) Small β_k, greater σ

Figure A4: $\Delta_i^S(x_{ij}|\lambda)$ where λ=(β, γ, σ) and i, j denote coordinates of a point in (lnK/L,α) space

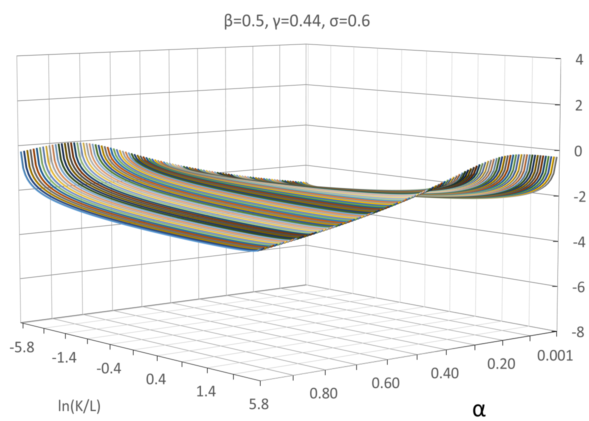
Source: Authors' calculations, see Appendix Section A.5.



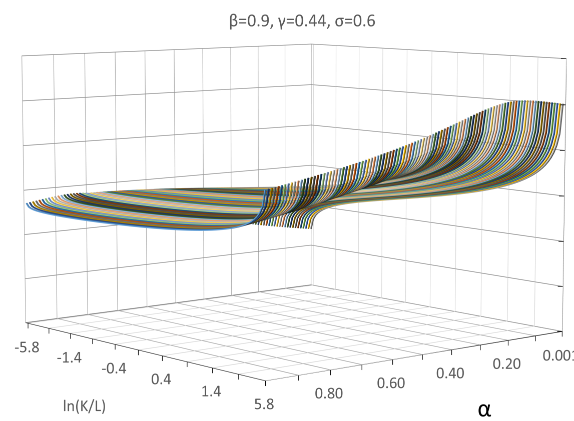
(a) Industry with equal CD elasticities



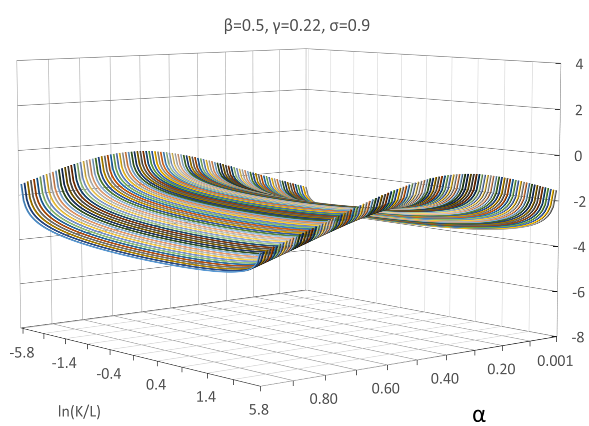
(b) Large β_k



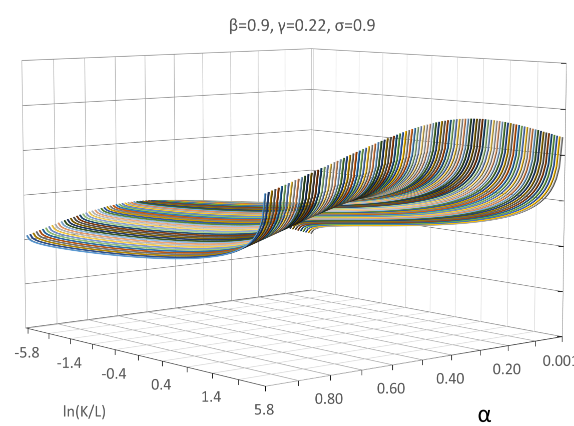
(c) Greater γ



(d) Large β_k, greater γ



(e) Greater σ



(f) Large β_k, greater σ

Figure A5: $\Delta_i^S(x_{ij}|\lambda)$ where λ=(β,γ,σ) and i,j denote coordinates of a point in (lnK/L,α) space

Source: Authors' calculations, see Appendix Section A.5.