Automation, Labor Share, and Productivity: Plant-Level Evidence from U.S. Manufacturing*

Emin Dinlersoz†
U.S. Census Bureau

Zoltan Wolf‡
New Light Technologies

May 3, 2019

Abstract

This paper provides plant-level evidence on the relationship between automation, factor usage, and productivity based on the U.S. Census Bureau's Survey of Manufacturing Technology. More automated plants exhibit lower production labor share, higher capital share, and higher labor productivity. They also experience greater long-term labor share declines and production labor productivity growth. Motivated by these patterns, a model of production with technology choice is estimated. The estimates indicate that more automated plants have higher total factor productivity and place less weight on production labor relative to capital, resulting in lower share and higher productivity for production labor.

JEL Codes: D24, O33, J30, L60
Keywords: automation, technology choice, total factor productivity, capital-labor substitution, labor share, CES production function, robots

---

*Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau or New Light Technologies. All remaining errors are our own. All results have been reviewed to ensure that no confidential information is disclosed.

†Center for Economic Studies, U.S. Census Bureau, 4600 Silver Hill Road, Suitland, MD 20746. E-mail: emin.m.dinlersoz@census.gov

‡New Light Technologies, 1440 G Street NW, Washington, DC 20005. E-mail: zoltan.wolf@census.gov

1 Introduction

The diffusion of automation is believed to be one of the fundamental drivers of the decline in employment and labor share, and the surge in output and productivity in U.S. manufacturing over the past decades. As robots increasingly take over the tasks performed by humans, the dependence on labor can recede further. These aggregate trends notwithstanding, micro evidence on the connection between automation, labor share, and productivity, has been scarce. This paper provides additional evidence on this connection using plant-level measures of automation from the U.S. Census Bureau's 1991 Survey of Manufacturing Technology (SMT). The SMT is ideal for studying patterns of automation-driven capital-labor substitution because it was designed to collect data from industries where capital better encapsulates advanced technologies that have the potential to replace labor.[1]

The stylized facts from the SMT, discussed in Section 2, indicate systematic relationships between automation and other plant-level outcomes. Specifically, labor's share in the value of shipments decreases as the degree of automation increases, driven by the negative relationship between production labor share and automation. In addition, more automated plants have a higher capital-production labor ratio, and a lower fraction of production workers, who are more productive and receive higher wages. Furthermore, plants with higher recent investment in automation experience larger declines in production labor share on a 5-to-10-year horizon. These findings are consistent with the subjective responses of plants in the SMT, which indicate that one of the most important benefits they receive from automation is reduction in labor costs.

Motivated by the stylized facts, we propose and estimate a model of CES production where plants choose thier production technology by adjusting the relative weight of capital and labor in response to relative price developments. A larger relative weight is interpreted as indication that the plant relies more

[1] Previous research used general capital measures, indicators of IT investment and use, and robot density as proxies for automation. In its broader definition, automation can include robots, machines with a pre-programmed computer software, metal-working lasers, optical inspection devices, automatic-guided vehicles, and many other technologies.

heavily on automation in production. This modeling choice is motivated by the fact that the large observed variation in the degree of automation is positively correlated with the relative weight. Relative factor prices play a central role in the model because they determine the relative weight. Modeling the plant's automation decision in this manner allows for alternative interpretations about the nature of automation, since any factor that affects the relative price will influence the plant's decision on automation. This approach is different from standard models of directed technical change, where the relative weight is typically governed by exogenous factor-augmenting shocks. It is based on the view that substitution may stem from a variety of factors including, but not limited to, such exogenous shocks. For instance, relative price variation may reflect differences in labor market frictions, labor quality and skills, financing constraints that limit technology adoption, unionization or the threat of it, the availability of outsourcing or offshoring, all of which may induce a plant to adjust the relative weight.

As in other models of CES production, the sensitivity of the relative weight to the relative price is characterized by the elasticity of substitution. This key technology parameter is estimated in the first step of the empirical analysis using plant-level information from the SMT. Given an estimate of the elasticity, the remaining parameters of the production function are determined following a version of the methodology described in Haltiwanger and Wolf (2018).[2]

Using factor weights or factor-augmenting shifters in CES production functions has a long tradition in the literature.[3] Exogenous factor-augmenting shocks are typically the primary source of heterogeneity in relative factor productivity. An immediate implication is that if there are other forces that generate heterogeneity in relative factor productivity then this assumption is arguably too restrictive. For example, if some production units actively invest in more automated capital and/or employ more skilled workers, then the

[2]The approach uses first-order conditions of profit maximization to determine the elasticity of variable factors. Quasi-fixed factor elasticities are estimated controlling for unobserved productivity differences using plant-level variation in advanced technologies investment available from the 1991 SMT.

[3]For recent examples, see Raval (2017) and Doraszelski and Jaumandreu (2018).

relative productivity could reflect differences in automation and labor composition. In addition, relative productivity and labor costs can be correlated in such cases. More generally, the relative weight and productivity of factors can be correlated with other plant-level outcomes. Although standard models with exogenous factor weights and productivity are able to generate correlations between these outcomes, the signs of these correlations are not always consistent with the accumulated evidence.[4]

Our approach is related to models of task-based automation, which emphasize the role of technological improvements in explaining the evolution of aggregate labor market outcomes.[5] These studies posit that the total output in the economy is a CES aggregate of individual task-level output. In contrast, the focus in this paper is specifically on the choices made by a production unit about the extent of automation used in production processes, motivated by the significant plant-level heterogeneity in the adoption and use of automation observed in the SMT.

The model of production in this paper is also related to models of endogenous productivity.[6] In these models, TFP depends on plant level decisions, such as R&D and exporting. It is less common to model factor-specific productivities or the relative factor weight as functions of plant actions. However, plant choices such as automation can be aimed at increasing the relative productivity of one factor. For instance, investing more in automation can increase labor productivity.

Earlier work has investigated capital-labor substitution using country and industry-level information on robots, automation, TFP, and relative input prices.[7] Given the more aggregate nature of these approaches, they are better

[4]Acemoglu and Restrepo (2018a) demonstrate the shortcomings of standard directed technological change models in capturing some hitherto documented effects of automation on labor outcomes. For example, they show that capital-augmenting technical change never reduces labor demand and increases labor share.

[5]See Acemoglu and Restrepo (2018a,b) for more details and about evidence that the model's implications are consistent with aggregate labor share and wage dynamics.

[6]See Ericson and Pakes (1995), Doraszelski and Jaumandreu (2018), Aw et al. (2011).

[7]Recent examples based on these variables are Elsby et al. (2013), Karabarbounis and Newman (2014), Graetz and Michaels (2018),Acemoglu and Restrepo (2017), and Autor and Salomons (2018).

suited to account for general equilibrium effects of automation. However, they are less informative about the connection between automation and production labor at the micro level. The analysis in this paper is intended to shed light on this connection by using direct measures of automation-related technologies at the plant level. The SMT contains information on the investment in, and use of, 17 such technologies, providing a level of detail unique in this literature.

The results in this paper also contribute to research on the decline in U.S. labor share.[8] Explanations of the decline include the diffusion of labor-saving technologies and automation brought about by the fall in the relative price of capital with respect to labor, import intensity and offshoring, the decline in unionization, or labor reallocation.[9] The last one of these has received a lot of attention with the rise of productive and large firms, and the associated increase in industry concentration of employment and sales. However, the microeconomic mechanisms through which these firms emerge have not been explored in detail. This paper offers additional evidence in this regard, which suggests that larger and more productive plants tend to be more automated and have lower production labor share, implying that increasing automation is one possible explanation behind the rise of such successful production units.

Part of the empirical analysis is related to previous literature that uses the SMT to analyze the connection between automation and plant-level outcomes. Most of the existing work relies on extensive measures or simple counts of technology types.[10] A number of papers look at the relationship between technology presence and plant life-cycle.[11]. Others explore the wage premia associated with technology use.[12] The SMT has also been used to study the

[8]In addition to the first two of the studies cited in the previous footnote, Lawrence (2015), Barkai (2016), Autor et al. (2017a,b) are recent examples.

[9]Autor et al. (2013) highlight the role international trade, while Elsby et al. (2013) consider the decline of unionization. Autor et al. (2017a,b) analyze the causes and consequences of labor reallocation.

[10]Beede and Young (1996) provide an extensive summary of this literature.

[11]Dunne (1994) explores the relationship between plant age and technology presence, while Doms et al. (1995) document that capital-intensive plants with advanced technology grow faster and are less likely to fail.

[12]Dunne and Schmitz Jr. (1995) find that establishments with more advanced technologies pay the highest wages and employ a higher fraction of non-production workers. Doms et al.

connection between labor productivity and technology.[13] Our work differs from these studies in two aspects. First, the analysis is based on intensive measures, such as the extent of investment in automation, and the degree to which a plant's operations depend on automation–instead of basic indicators of automation presence. In addition, the objective is to estimate a model of production in a way that accounts for the connection between input prices and automation, and at the same time controls for unobserved productivity differences. For the latter purpose, intensive indicators of investment in automation are more appropriate because they measure these differences more accurately.

2 Data

The main data source is the U.S. Census Bureau's 1991 version of the SMT, part of three separate surveys on technology use in manufacturing plants conducted in 1988, 1991, and 1993. It contains a rich set of measures on the automation-related technologies. Extensive measures on technology presence are available in the 1988 and 1993 versions, while intensive measures on technology use and investment are recorded in the 1991 version. The 1991 survey contains a stratified random sample of about 10,000 observations, representative of nearly 45,000 plants.

The 1991 SMT has data on 5 broadly defined manufacturing industries[14]: Fabricated Metal Products (SIC 34), Industrial Machinery and Equipment (SIC 35), Electronic and Other Electric Equipment (SIC 36), Transportation Equipment (SIC 37), and Instruments and Related Products (SIC 38). These industries were included in the survey because they rely more on automation-related technologies, which makes them ideal for the purposes of this paper.[15]

(1997) document that businesses with a greater number of advanced technologies have more educated workers, more managers and pay higher wages. They do not find, however, a significant correlation between skill upgrading and use of advanced technologies.

[13]SeeMcGuckin et al. (1998).

[14]These industries accounted for about 43% of manufacturing employment in 1991.

[15]The SMT was funded partly by defense agencies. During the developmental phase the Census Bureau consulted with government agencies, private industry and academic experts. The industries were selected because of the relatively high presence of defense contractors, which also tend to be more advanced in terms of technology. A number of studies in

The SMT industries are not substantially different from the rest of manufacturing in terms of average capital and production labor shares, see Table 1(a). Although the non-production labor share is lower and the capital-labor cost-ratio is higher, these indicators are also comparable to the rest of manufacturing. The fact that the SMT industries do not stand out in terms of standard measures suggests that these measures may not be informative by themselves about the underlying extent of automation in an industry.[16]

2.1 Technology Questions

The SMT contains information on the adoption and use of 17 individual technologies, see Table 1(b), some of which are directly aimed at automation, whereas others can facilitate or support automation.[17] The current analysis views all of the technologies as part of the automation in the plant because they all have the potential to replace production workers.

The 1991 version of the SMT records information about use and investment, the 1988 and 1993 versions contain information about the presence of technologies. The 1991 SMT has information only on 4 broad technology groups, which contain 17 detailed technologies which were the focus of the 1988 and 1991 surveys. While the SMT pertains to an earlier period, many of the technologies had already diffused to a large extent by the time of the survey –see Table 1(b). While robots were less common around the time of the survey, other automation-related technologies were relatively wide-spread. The similar relative diffusion rates in 1988 and 1993 suggest that the 1991 diffusion rates are likely similar.

engineering economics support this view, see, e.g., Kelley and Watkins (1998). This finding echoes in the SMT: plants that indicate production to military specs have on average higher technology use and investment.

[16]Details about the evolution of these industry aggregates can be found in Appendix A.1. The industries exhibit similar trends, and the decline in labor share is consistent with the well-documented decline for overall manufacturing.

[17]The former group includes Robots, Automated Storage and Retrieval Systems, Automated Guided Vehicle Systems, and Automated Sensor Based Inspection/Testing Equipment, the latter group includes Computer Aided Design/Engineering, Computer Aided Manufacturing, Local Area Networks. Exploring which technologies matter most for replacing labor is important but is beyond the scope of the paper and is left for future work.

The key data source for this paper is the 1991 version, because intensive measures, such as the fraction of automated operations, and the dollar-value of investment in automation, more accurately capture cross-establishment differences in the extent of automation than the simple technology counts. The technology indicators used in this paper are based on 4 survey questions about current and planned use of, and investment in, 4 broad technology groups, see Table 2. Responses are recoded into numerical categories, implying that the interpretation of exact quantitative differences across categories is difficult. Nevertheless, these categories capture some of the cross-plant variation in technology use and investment, which is the key variation used to control for unobserved productivity differences during production function estimation.

2.2 Technology Index

In the empirical analysis, the degree of automation is measured using information about current, as well as planned, investment and use pertaining to each of the four technology groups, described in Table 1(b).[18] In particular, plant-level responses to the questions in Table 2 are aggregated into a plant-specific average. The resulting average index is continuous between 0 and 5, where zero indicates no automation, a positive value represents the average of the extent to which the plant uses and invests in automation.[19]

2.3 Input and Output Measures

Plant-level measures of output and inputs were obtained from the 1991 Annual Survey of Manufactures (ASM). Since the sampling frame of the ASM is different from the SMT, some of the records could not be matched. In such cases, the 1990 ASM and 1992 Census of Manufacturers (CMF) were used

[18] Incorporating the response to the question on the anticipated cost of future acquisitions makes little difference in our results and conlcusions.

[19] An alternative measure based only on the investment question (2) in Table 1(b) is also considered. To prevent confusion, whenever both measures are used in the empirical analysis, the main technology index is labeled as "Technology Index I", whereas the investment-based one is labeled as "Technology Index II".

to supplement the output and input measures.[20] Output is measured as the deflated value of total value of shipments. Production labor input is measured by production worker hours, non-production worker input is calculated as the product of production worker hours and the ratio of non-production wage bill to production wage bill. Intermediate inputs are calculated as the sum of cost of parts, contracted work and goods resold. The energy input is composed of costs of electricity and fuel. Capital stock measures are based on a version of the Perpetual Inventory Method that generates current capital by summing the depreciated stock and current investment. The initial capital stock is a deflated book value that is taken from the ASM-CMF.[21]

A separate panel of plants is also used to check the robustness of the elasticity of substitution estimates. This panel is constructed using all ASM plants from SMT industries, between 1987 and 1996. For the analysis of the relationship between the degree of automation and the evolution of labor share and labor productivity, the plants in the 1991 SMT that survive and appear in the 1997 and 2002 CMF were identified using the U.S. Census Bureau's Longitudinal Business Database (LBD). The surviving plants are used to study the evolution of labor share within the next 5 to 10 years as a function of the degree of automation in 1991.[22]

3 Stylized Facts

3.1 Labor Share and Automation

Based on plants' own assessments in the 1991 SMT, quality improvement and labor cost reduction are the two most important benefits from automation, see Figure 1 for more details. In other words, plants recognize the importance of input prices in automation investment, which is especially relevant for the purpose of the current analysis. Other benefits include, in decreasing order of

[20] If an SMT-plant cannot be matched to the 1991 ASM, it is matched to the 1992 CMF. Otherwise, the plant is matched to the 1990 ASM.

[21] More details on these data can be found in Foster et al. (2017).

[22] Survival bias is also accounted for in this analysis for robustness.

importance, the increase in the flexibility of production, lead-time reduction, and marketing advantage.

In order to get a more accurate picture of input use, Figure 2(a) plots plants' labor share in total value of shipments as a function of the technology index discussed in Section 2.2. Labor share is lower for more automated establishments: it drops from 29% to 24% as the technology index increases from 0 to 5. The decline is statistically significant for much of the index range, and is driven by production labor.[23] The share of production labor drops nearly by half moving from the lowest to highest index levels (Figure 2(b)), whereas non-production labor's share increases only slightly (Figure 2(c)). More automated plants also tend to have a lower fraction of their workforce engaged in production (Figure 2(d)). This fraction drops from 70% to 50% across the index range. In addition, capital's share in the value of shipments increases with the technology index (Figure 3(a)). As a result, the ratio of capital share to labor share, and the capital-labor ratio also increase with the technology index (Figures 3(b)-3(c)).[24] The technology index is also correlated with other plant level outcomes. Labor productivity increases with the index, especially for production labor (Figure 4), a finding robust to alternative ways to measure labor productivity (Figure A5 in Appendix A.5). Average wage (total wage bill divided by the number of workers) also increases with the index for both types of labor (Figure A4).

These relationships are robust to a broad set of controls such as plant size and age, unionization of production workers, foreign-ownership, military production, or whether or not a plant exports. Tables 3(a)-3(c) show the estimated coefficients of the technology index conditional on these controls.[25]

[23]The confidence intervals get larger at the high end of the technology spectrum because sample size decreases and because of the one-sided nature of the kernel smoothing near the end of the sample.

[24]The patterns in Figures 2 and 3 continue to hold if plant value added is used instead of revenues, when industry effects are netted out, or when other plant characteristics are controlled for.

[25]For robustness, these tables also feature the alternative technology index (Technology Index II) that only includes the average investment indicator across the four technology groups based on survey question 2 in Table 2.

The results indicate that, controlling for observables, a 1% increase in technology index is associated with a 0.04-0.08% decline in production labor share, 0.12-0.14% increase in production labor productivity, and 0.08-0.09% increase in average production worker wage across plants. In contrast, the technology index does not seem to be related to non-production labor share, while average wage and labor productivity of non-production workers both increase with the technology index.[26] 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.[27]

3.2 Change in Labor Share and Automation

A full dynamic analysis of the effects of changes in automation is not possible because the technology index can be constructed in 1991 only.[28] However, we can analyze establishment-level outcomes as a function of the 1991 automation status, which allows suggestive conclusions about dynamic effects because many automation-related technologies are likely to remain in place for longer time periods. To explore the connection between automation and change in labor share, the following specification is estimated

$$\Delta y_i = b_o + b_I I_i + b_E \Delta e_i + b_X X_i + \varepsilon_i, \tag{1}$$

where Δy_i denotes the log difference in the labor share or labor productivity between 1991 and 1997, or between 1991 and 2002. Δe_i denotes the log difference in the plant's total employment over the same horizon, it serves to control

[26] 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.

[27] 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.

[28] An additional argument for using the 1991 SMT is that there is significant attrition between the 1988 and 1993 surveys, and the technology indicators in these surveys do not reflect investment or usage intensity. Prior research with these two surveys also indicate some potential recall bias.

for growth-related heterogeneity.[29] I_i denotes the 1991-value of the technology index, and X_i contains plant-level controls and industry effects.[30] The results in Tables 4(a) and 4(b) indicate that plants that were more automated in 1991 tend to experience lower production labor share growth and higher production labor productivity growth over the next 5 to 10 years. These results, taken together with those in the previous section, confirm that the correlation between automation, labor share and labor productivity is statistically and economically significant.[31]

4 The Model

This section describes a model of plant-level production that is consistent with the stylized facts discussed above. A key feature of the model is technology choice: the plant adjusts the relative weight of capital and production labor in response to changes in their relative price.

Plant i generates output according to the production function

$$Q_i = \theta_i L_{ni}^{\beta_1} M_i^{\beta_2} E_i^{\beta_3} [\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}]^{\gamma/\rho}, \tag{2}$$

where θ denotes Hicks-neutral productivity. For simplicity, freely variable inputs non-production labor (L_n), intermediate inputs (M), and energy (E) are combined into a Cobb-Douglas (CD) aggregate, with elasticities β_j.[32] The focus is on the capital (K) and production labor (L_p), which are aggregated into a constant-elasticity-of-substitution (CES) composite, $T_i = [\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}]^{1/\rho}$, where σ denotes the elasticity of substitution between L_p and K, and $\sigma = (1-\rho)^{-1}$.

[29] Growing plants hire more employees and therefore their labor share is expected to rise.

[30] Because Δy_i is observed only for plants surviving till 1997 (or 2002), a Heckman two-step estimation is also implemented to account for the bias introduced due to this selection.

[31] Controlling for survival bias using a Heckman correction confirms these conclusions: Tables A2 and A3 in Appendix A.5 show qualitatively similar results. Results are also stronger when value added is used to measure production labor share and productivity.

[32] Figure 2(c) indicates no significant correlation between non-production labor share and the degree of automation, suggesting that the simplification is reasonable.

K and L_p are assumed to be quasi-fixed. This assumption is justified by existing micro evidence that both capital and labor adjustment are subject to non-linearities – see, e.g., Caballero et al. (1997), Cooper and Haltiwanger (2006) and Bloom (2009). In addition, a union contract for production workers exists in nearly 20% of the 1991 SMT plants, which is an indication that production labor is subject to similar rigidities – see Dinlersoz et al. (2017). Assuming that L_n is a freely variable input simplifies the analysis and helps maintain focus on the connection between L_p and K.[33]

The relative weight of K and L_p in T is determined by $\alpha_i \in (0, 1)$, which is referred to as the degree of automation in the plant. As discussed in more detail in Section 4.1, the first order conditions of cost minimization imply that in equilibrium the plant adjusts the relative weight $\alpha_i/(1 - \alpha_i)$ in response to changes in the relative price of K and L_p, conditional on σ. This is a deviation from earlier work on CES production because the relative weight is generally assumed to be exogenously given, and may even be a constant across establishments, see, e.g. Lawrence (2015) and Raval (2017).

The decision variable α_i can be interpreted in a number of ways. A higher α_i in standard models represents a technology that augments the productivity of K more relative to L_p.[34] Alternatively, given the positive relationship between K and automation in the SMT, a higher α_i can also indicate that the establishment has more automation-related capital. What underlies the latter interpretation is the plant's decision on α_i in response to a change in the relative price. Specification (2) can also be interpreted as a reduced-form, plant-level representation of a task-based production model, analogous to the approach in Acemoglu and Restrepo (2018a,b).[35] Suppose that plant-level production consists of a set of individual operations indexed on the unit interval [0,1], each of which can be carried out using either capital or production

[33]It is possible to generalize the analysis using nested CES specifications that allow varying substitutability between both labor inputs and capital, as well as energy and materials.

[34]To see this, rewrite T_i in (2) as $[(A_i K_i)^\rho + (B_i L_{pi})^\rho]^{\gamma/\rho}$ and set $A_i = \alpha_i^{2(1-\rho)/\rho}$ and $B_i = (1 - \alpha_i)^{2(1-\rho)/\rho}$.

[35]However, those models do not pertain to a production unit, and task-level output is aggregated to the economy-level.

labor. The plant chooses α_i, the fraction of operations it wants to automate, and the amount of K or L_p conditional on α_i. In this context, a higher α_i can be interpreted as a higher degree of automation.[36]

It is straightforward to generalize the weights in the composite input T_i by rewriting it as T_i=$[\alpha_i^{\zeta/\sigma} K_i^{\rho} + (1-\alpha_i)^{\zeta/\sigma} L_{pi}^{\rho}]^{1/\rho}$, where $0 < \zeta$ is a parameter. Specification (2) then corresponds to the case of $\zeta = 2$. Although choosing $\zeta = 2$ restricts the parameter space, the restriction is not inconsistent with the properties of the SMT, while it delivers analytical convenience and also limits the number of parameters to estimate.[37]

Note that in a fully specified model, Hicks-neutrality implies that α_i affects only T_i, so the production function in (2) does not restrict the relationship between productivity and other plant characteristics. However, if (2) is not fully specified, θ_i and other plant characteristics may be correlated.[38]

4.1 The Plant's Problem

Throughout this section, plants are assumed to be price takers in input markets – a standard assumption in the empirical productivity literature. In the first part, price taking behavior is also assumed for output markets.

4.1.1 Exogenous Output Prices

Plants produce a homogenous good with a fixed price, normalized to one. All factor prices, denoted by w_{ji}, are allowed to vary across plants, as opposed

[36]Here, no formal equivalence to task-based models is claimed. Although in principle such a model could be posited at the plant-level, this approach would require additional structure and therefore is left for future work.

[37]Noting that CD is a special case of (2) with $\sigma = 1$, if the true data generating process in the SMT can be described by (2), then we should expect the σ-estimates to be less than one. We show in Section 5.1 that the value of ζ has implications for the estimated value of σ, and as we will see in Section 6, $\zeta = 2$ yields elasticity estimates that are consistent with these expectations, given the properties of the SMT.

[38]For instance, it may be the case that θ_i=$\theta(\alpha_i)$, where α_i directly affects Q_i in addition to its effect on T_i. Such effects are plausible if adopting labor-saving technologies also results in more flexible production, and improves overall coordination and monitoring of production processes, resulting in higher Hicks-neutral productivity.

to the typical assumption that they are constant. The assumption of heterogenous input prices is justified if, for instance, there are differences across plants in terms of the quality of their inputs. One example would be the case in point where plant-level capital stocks differ in the extent to which they contain automation-related technologies.[39] Under these assumptions, the first-order conditions of cost minimization can be written as[40]

$$K_i/L_{pi} = (w_{pi}/w_{ki})^{2-\sigma} \tag{3}$$

$$\alpha_i/(1-\alpha_i) = (w_{pi}/w_{ki})^{1-\sigma}\,. \tag{4}$$

These expressions highlight the key mechanism of the model: both the capital-labor ratio and the relative weight are tied to relative input price variation and the nature of these relationships is fully determined by σ. It follows from equations (3) and (4) that K_i/L_{pi} and $\alpha_i/(1-\alpha_i)$ are increasing and convex in w_{pi}/w_{ki}, as long as $\sigma < 1$. That is, an increase in the relative price of L_p induces the plant to substitute K for L_p by increasing α_i, and more so for higher values of α_i. Equations (3) and (4) also imply that the capital-labor ratio can be written as a function of the relative weight:

$$K_i/L_{pi} = (\alpha_i/(1-\alpha_i))^{\frac{2-\sigma}{1-\sigma}}\,, \tag{5}$$

and that this relationship is also convex because $1 < (2-\sigma)/(1-\sigma)$ as long as $\sigma < 1$. Equation (5) formalizes the basic idea that automation increases the relative weight and the capital-labor ratio.

The implications of equations (3)-(5) are consistent with the observed relationships in the SMT, see the discussion of Figures 3(b)-3(c) in Section 3. Importantly, the implications are different from alternative models. To be specific, these relationships are not convex in the case of a standard CES model with $\sigma < 1$, or under Cobb-Douglas production.[41]

[39] Input price variation can also be a result of differences across locations, amenities, agglomeration economies, and costs of mobility and adjustment may imply persistent differences in the price of labor and capital.

[40] See Appendix A.2 for more details.

[41] In the standard CES model with exogenous relative weight, i.e. $[\alpha_i K_i^{\rho} + (1-\alpha_i) L_{pi}^{\rho}]^{\gamma/\rho}$,

It is useful to describe the properties of shares of input expenditures because they can be used to develop estimators. By combining equations (3)-(4) it can be shown that L_{pi}'s optimal weight in production equals its share in the cost of T_i:

$$1 - \alpha_i = w_{pi}L_{pi}/(w_{ki}K_i + w_{pi}L_{pi}). \tag{6}$$

Using the fact that the revenue share of T_i equals the elasticity, γ, of Q_i with respect to T_i, the revenue shares of L_{pi} and $(L_{pi} + L_{ni})$ can be written as

$$w_{pi}L_{pi}/Q_i = \gamma(1 - \alpha_i) \tag{7}$$

$$(w_{pi}L_{pi} + w_{ni}L_{ni})/Q_i = \beta_1 + \gamma(1 - \alpha_i). \tag{8}$$

Equations (6), (7) and (8) indicate that all labor usage indicators are decreasing in α_i but the revenue shares' sensitivity to α_i depends on γ.[42] The cost share of the jth variable input can be written as $cs_j = \frac{\beta_j}{\sum_j \beta_j + \gamma}$, and the share of T_i in total costs is given by $cs_{Ki} + cs_{L_{pi}} = \frac{\gamma}{\sum_j \beta_j + \gamma} \times c_i$, where $c_i = \frac{\alpha_i^{2/\sigma} K_i^{\rho-1} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho-1}}{\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}} < 1$ if $\sigma < 1$.[43] One difference relative to CD technology is that imposing constant returns to scale (CRS) is not a sufficient condition for identification. Even though variable input elasticities are identified by cost shares under CRS, the share of T_i in total costs underestimates the contribution of γ to returns-to-scale, irrespective of its value.

If returns-to-scale is unknown, the implications of profit maximization can be used to recover factor elasticities. The first-order condition from profit maximization imply that factor elasticities can be written as $\beta_j = w_{ji}X_{ji}/Q_i$, and $\gamma = w_{ki}K_i/Q_i + w_{pi}L_{pi}/Q_i$, which show that under exogenous prices and unknown returns-to-scale, the factor elasticities of both freely variable inputs and the composite input are identified by revenue shares of input expenditures.[44]

we have $K_i/L_{pi} = (\alpha_i/(1-\alpha_i) \times w_{pi}/w_{ki})^{\sigma}$, so when $\sigma \in (0,1)$, K_i/L_{pi} is concave in both ratios. CD production implies $K_i/L_{pi} = \alpha_i/(1-\alpha_i) \times w_{pi}/w_{ki}$, so K_i/L_{pi} is linear in the ratios.

[42] When γ=1, all three shares decline one-for-one with α_i. If $\gamma < 1$, the revenue shares' rate of decline is smaller, and the relationship is reversed when $\gamma > 1$.

[43] This statement holds if $1 < L_{pi}, K_i$, which is the case in yearly data on these variables.

[44] See Appendix A.2 for details.

4.1.2 Isoelastic Residual Demand

An alternative to fixed output prices is to postulate that the plant's residual demand is isoelastic, a commonly used approach in the literature.[45] Under this assumption, the inverse residual demand function can be written as $P_i = P\left(Q/Q_i\right)^{1-\kappa}\xi_i$ with $0 < \kappa < 1$, where P and Q denote aggregate variables and ξ_i is an idiosyncratic demand shifter. The results of cost minimization are robust to alternative assumptions about demand. The conclusions of profit maximization are different because under downward sloping demand conditions marginal revenue products are smaller than marginal products. To see this, let $R_i{=}P_iQ_i$ denote plant-level revenues, and write the first order conditions for the jth variable input and the quasi-fixed inputs as

$$w_{ji}X_{ji}/R_i = \kappa\beta_j \tag{9}$$

$$(w_{ki}K_i + w_{pi}L_{pi})/R_i = \kappa\gamma, \tag{10}$$

where (10) combines conditions for K and L_p. The implications of (9)-(10) for the relationship between w_{pi}/w_{ki}, K_i/L_{pi} and $\alpha_i/(1-\alpha_i)$ are the same as in Section 4.1.1 because relative input allocations are not affected by ξ_i.

An important difference relative to Section 4.1.1 is that the revenue share of input expenditures now depend on both β_j and κ. In principle, information on output prices could be used control for output price variation during estimation, which in turn would allow the identification of factor elasticities. However, output prices in SMT are recorded as a categorical variable and preliminary analysis indicates that this variable has no additional explanatory power conditional on the variables already included in the analysis, potentially due to the relatively coarse price categories used in the survey.[46]

[45] For Recent examples, see De Loecker (2011), Bartelsman et al. (2013), Foster et al. (2016, 2017), and Haltiwanger and Wolf (2018).

[46] The categorical price variable measures average price for the products of a plant and is available in the 1988 and 1993 SMT only. The price information was merged for the set of plants in the 1991 SMT that are in the union of the 1988 and 1993 SMT plants.

5 Estimation

The estimation strategy uses some of the model properties discussed in Section 4. The elasticity of substitution is estimated using a log-linearized version of (3). The remaining parameters in (2) are estimated conditional on σ, using a modified version of the method described in Haltiwanger and Wolf (2018).

5.1 Elasticity of Substitution

Log-linearizing equation (3) yields $l_{pi} = (\sigma - 2)\ln w_{pi} - (\sigma - 2)\ln w_{ki} + k_i + \varepsilon_i$, where ε_i is an *iid* error. Given data on l_{pi}, k_i, and their prices, this equation can be estimated by running the regression

$$l_{pi} = \delta_1 \ln w_{pi} + \delta_2 \ln w_{ki} + \delta_3 k_i + u_i. \tag{11}$$

Since w_{pi} is obtained by dividing production labor costs by production worker hours, OLS estimates of δ_1 are affected by division-bias. This issue is addressed using geographic variation in w_{pi} as instruments, where the instrument is defined as a county-specific average manufacturing wage indicator, calculated using plant-level information. This approach is similar to the method used by Raval (2017). The rental price is also unobserved in both the SMT and the ASM. The plant-level measure of w_{ki} is calculated by combining industry-specific rental prices and plant-level capital cost measures, which provide $w_{ki}K_i$ and K_i but not w_{ki}, implying that δ_2 is not identified.[47] Under the assumption that the rental price is plant-specific, its effect can be accounted for by a plant-fixed effect, in which case δ_3 is identified in the first-differenced version of (11). This strategy is justified when plant-level capital prices are persistent, for instance when they follow a random walk.[48]

Before moving on to discussing the estimation of factor elasticities, it is useful to reiterate a point that concerns restricting the parameter space. Con-

[47] See Foster et al. (2016) for more details on capital measures.

[48] Equation (11) has no explicit dynamic elements in the econometric sense despite the fact that the plant's optimal decision on both capital and production labor have dynamic implications, see Section 4. This is because the main purpose of (11) is to estimate long-run patterns of capital-labor substitution.

ditional on an estimate of δ_1, $\widehat{\sigma}$ is calculate as $\widehat{\sigma}=\widehat{\delta}_1+2$ because we chose $\zeta=2$ in equation (2). As mentioned in Section 4, this choice is motivated by analytical convenience and the fact that it supports σ-estimates that are consistent with the implications of the model and the observed patterns in the SMT. In particular, the stylized facts suggest that K and L_p are less substitutable than what a CD specification would imply, and hence one expects $\widehat{\sigma}<1$. Choosing $\zeta=2$ allows for this possibility and also helps maintain analytical tractability.

5.2 Factor Elasticities

The estimation strategy for factor elasticities is based on earlier results in the empirical productivity literature, but also deviates from standard approaches in order to better make use of the features of the SMT. The 1991 SMT records categorical responses on how much the plant invested in four technology types in the previous three years, see question 2 in Table 2. Although the variables are categorical, they provide direct information on cross-plant differences in automation investment. The responses are aggregated into a plant-level indicator of technology investment, which is then used as a proxy to control for unobserved productivity differences during estimation. This proxy is a distinguishing feature relative to earlier studies because those relied on general investment to control for unobserved productivity differences.[49]

In addition to the unique proxy, the estimation strategy deviates from the standard proxy-based approaches in two other respects. First, it abstracts from selection because the SMT has limited information on investment history. Second, it follows the methodology described in Haltiwanger and Wolf (2018) to non-parametrically estimate the elasticities of freely variable inputs. Under the assumption of downward-sloping demand, the revenue shares of variable input expenditures depend only on the corresponding factor elasticity and the demand parameter, implying that revenue shares can be used to identify elasticities without projecting revenue variation on proxies, state variables

[49]The idea of accounting for unobserved productivity differences during estimation by using firm-level proxies is discussed in Olley and Pakes (1996) and Levinsohn and Petrin (2003).

and variable inputs.[50] This property is useful because Gandhi et al. (2016) show that the identification of intermediate input elasticities is problematic when using intermediate inputs as a proxy. Given estimates of variable input elasticities, Haltiwanger and Wolf (2018) propose to net out the contribution of variable input expenditures from revenue variation, and use this net variation to estimate the remaining coefficients. The main difference relative to Haltiwanger and Wolf (2018) is how the net variation is used to determine the remaining coefficients, since their approach considers Cobb-Douglas technology. A more detailed description of the estimation strategy can be found in Appendix A.3.

6 Results

6.1 Estimates of the Elasticity of Substitution

The estimates of σ vary between 0.38 and 0.71 depending on the methods used, see Table 5(a). These estimates are consistent with what recent work found using similar data from the U.S. Census Bureau.[51] The baseline $\widehat{\sigma}$, shown in column 1 of Table 5(a), is determined by a cross-sectional IV. Column 2 contains the results without k_i as an explanatory variable. This approach may be justified if w_{pi} and L_{pi} are measured with less noise than capital, which may be the case for ASM.[52] If capital is measured with error then it is a priori unclear whether including k_i in (11) is useful. The similarity of the σ estimates is reassuring and at least suggests that production labor data alone is informative for substitution patterns. Columns 3 and 4 show additional robustness checks, where (11) is estimated using ASM data from SMT industries between 1987 and 1996. Instead of using geographic wage variation as an instrument, these calculations are based on lagged differences

[50] See Section 4.1.2 for more details.

[51] For instance, Raval (2017) estimates a plant-level elasticity of substitution between labor and capital in the range 0.3-0.5, and Oberfield and Raval (2014) report estimates between 0.4 and 0.7.

[52] This survey collects data on book-value capital, which is then converted into market values using data on depreciation and various deflators available at the industry level only.

of plant-level wages as instruments in a GMM framework, see Arellano and Bond (1991). The GMM estimator yields similar $\widehat{\sigma}$s.

As previously discussed, $\sigma < 1$ implies that capital and production labor is less substitutable than a Cobb-Douglas (CD) specification would imply.[53] The lower estimated substitutability is plausible considering that the plants in the SMT typically utilize automation, which may give rise to complementarities between K and L_p. Given $\sigma < 1$, the relative weight $\frac{\alpha_i}{1-\alpha_i}$ is increasing and convex in the relative price.

6.2 Estimates of Factor Elasticities

Table 5(b) shows estimated factor elasticities conditional on the baseline $\widehat{\sigma}$. As the column labeled "$\widehat{\gamma}$" indicates, all reviewed methods yield comparable γ estimates suggesting that the contribution of the composite input to returns-to-scale is between 0.17 and 0.25, depending on whether it is determined using simple plant-level averages of the capital and production labor expenditures in revenues (row 1) or projection-based methods (rows 2-3). The sum of factor elasticities is less than one. However, without data on prices and quantities these point estimates are revenue elasticities implying that they can be considered as lower bounds for factor elasticities – see Section 4.1.2 for more details.

6.3 Properties of Productivity Estimates

This section investigates the properties of the implied total factor productivity (TFP) distributions. For robustness, two commonly used CD productivity measures are evaluated against three CES measures. The first CD measure, denoted by CD_{CRS}, is derived under the assumption of constant returns-to-scale. The second one, labeled as CD_{NCRS}, is calculated under the assumption of non-constant returns-to-scale. To simplify discussion, we maintain the assumption of homogenous products, and price taking behavior.[54] The CES productivity indices correspond to the three specifications shown in Table 5(b). The first of these, denoted by CES_{FOC}, is based on $\widehat{\gamma}$ obtained as a plant-level average

[53] In other words, CES isoquants have more curvature than CD isoquants.

[54] Analyzing the role of demand is deferred to future work.

of condition (10) under the assumption technology choice. The second CES measure, labeled as CES_{EN}, is a variant of the first one, in the sense that γ is estimated using nonlinear least squares. The third one, labeled as CES_{EX}, is similar to the second one except that α_i=1/2 is imposed as a reference point for a case where α_i is an exogenously given constant across plants.

The descriptive statistics, shown in Table 6, suggest that the shape of the TFP distribution implied by CES specifications is generally different from those under CD specifications. Although all five distributions have negative skew indicating that the left tails are longer, there are differences in how dispersed and slender they are. CD yields more observations around the mode and in the tails, indicated by higher kurtosis and lower dispersion. Bivariate correlations in Table 6(b) echo these differences: while the association among alternatives derived from the same technology is strong, the correlation between CD and CES residuals is significantly less than one.

In light of these findings, it is natural to ask whether the TFP distributions implied by alternative technologies are systematically different. Given that the SMT collected data from industries where automation is a likely substitute for labor, the distinction between CD and CES is expected to be relevant. The p-values of the Kolmogorov-Smirnov (KS) tests in Table 6(c) confirm this conjecture: the CD and CES measures are significantly different at usual levels of significance. Interestingly, the assumption of endogenous versus exogenous technology also matters. In addition, the way elasticities are calculated under CD also matters. The only pair for which the null of equivalence cannot be rejected is the one where γ is estimated using different methods.

The KS test is useful to test whether two distributions can be considered different in the statistical sense. However, it is not informative about possible sources of the difference. In order to shed some light on the nature of the differences discussed above, Appendix A.4 provides a detailed decomposition in which the difference between CD_{CRS} and CES_{EN} is parsed into a term that is due to differences in the functional form, and additional components that can be attributed to estimation error. The contribution by the difference in functional form can be interpreted as an estimate of the specification error in

the population if the true data generating process is CES_{EN}. The difference can be written as

$$\Delta_i = \frac{\widehat{\gamma}}{\widehat{\rho}} \ln \left[\alpha_i^{2/\widehat{\sigma}} K_i^{\widehat{\rho}} + (1-\alpha_i)^{2/\widehat{\sigma}} L_{pi}^{\widehat{\rho}} \right] - \left(\widehat{\beta}_k \ln K_i + \widehat{\beta}_l \ln L_{pi} \right). \tag{12}$$

It is a useful metric because it helps understand why the KS test rejects the null of equivalence. Interpreted as a sample statistic, Δ_i accounts for all the specification error on condition that the estimation error is the same under the two specifications. Appendix A.4 explores the properties of Δ_i in more detail. It turns out that Δ_i is negative for the majority of plants, meaning that the CD_{CRS} input index in the second term in (12) is systematically higher than the CES_{EN} input index in the first term of (12). In other words, CD_{CRS} tends to underestimate productivity if the true underlying productivity is CES_{EN}. The results also indicate that the extent of this error tends to be higher for plants with higher capital-production labor ratio and more automation. These findings imply that accounting for capital-labor substitution patterns in productivity estimation is potentially important for correctly measuring TFP.

6.4 Automation, Productivity, and Labor Share

In order to shed light on the relationship between total factor productivity and automation, Table 7(a) reports the results from regressing the revenue share of L_p on the estimated TFP (CES_{EN}) and technology index, controlling for other plant characteristics. The main conclusion is that the revenue share of L_p is lower in more productive and automated plants and that productivity is more negatively correlated with L_p share than with technology.[55]

Table 7(b) repeats the previous analysis using $1-\alpha_i$ as the dependent variable, see equation (6). The results indicate that technology is more strongly associated with this measure.[56] Interestingly, $1-\alpha_i$ is positively associated

[55]The results are similar if value-added is used instead of total value of shipments in defining labor share.

[56]Comparing (6) and (7) shows that if $\gamma < 1$ then $\gamma(1-\alpha_i) < (1-\alpha_i)$, and therefore a given change in α_i should imply a smaller decline in L_p cost's share in revenue than in composite input expenditures.

with productivity and negatively associated with plant size, which are the opposite of the estimated signs in Table 7(a), suggesting that choice of the labor share measure (revenue- or cost-based) matters for subsequent analyses.

If the model in Section 4 is fully specified, then α_i and θ_i are not correlated. However, a systematic relationship may emerge between the two measures if not all factors of production are accounted for during estimation. In other words, automation may be correlated with productivity in the presence of relevant unobserved heterogeneity. In order to assess the presence of such factors, the relationship between technology indices, productivity and other observables is assessed in a regression framework. Table 7(c) contains the estimated coefficients, which indicate that more productive and larger plants tend to be also more automated.[57] It is important to note that the relationships change with CD_{CRS}. Table A4 in Appendix A.5 shows the details, here we only point out that CD_{CRS} is more negatively associated with the share of production labor in revenue and shows no significant association with the share of production labor in composite input expenditures. These findings indicate that the specification of the production function matters. One implication is that an incorrectly specified production function can have important consequences for heterogenous agent models that aim to capture specific relationships regarding productivity, automation, and capital-labor substitution.

7 Concluding Remarks

There is a growing body of theoretical and empirical work on the aggregate effects of automation on employment, labor share, and productivity. However, micro-evidence on the connection between these variables has been limited, due mainly to lack of detailed measures on automation. This paper provides additional evidence on these relationships using plant-level information on au-

[57] It is a common finding in the productivity literature that the productivity residual is correlated with other plant characteristics. It indicates the presence of factors that this simple model does not account for. For instance, automation may enhance managerial ability, inventory management, or coordination in factory floor, which are factors that are not fully captured by standard measures of input usage.

tomation from the U.S. Census Bureau's 1991 Survey of Manufacturing Technology (SMT). The stylized facts from the SMT indicate that plants with greater use of, and investment in, automation have higher capital share and lower production labor share in revenue. More automated plants have relatively lower fraction of production workers, exhibit higher labor productivity, and pay higher wages.

Motivated by these facts, a model of plant-level production is proposed and estimated. A key property of the model is that the plant chooses the relative weight of capital and labor in a CES composite. The relative weight can be interpreted as the plant's choice of the degree of automation. This modeling choice is a deviation from standard models of biased technical change where the relative weight of inputs is determined by exogenous factor-augmenting shocks. Endogenous technology choice reflects the view that substitution between capital and labor may stem from any shock that alters the relative price of these inputs. Therefore, the model supports alternative interpretations about the underlying reasons for capital-labor substitution.

The model's predictions are consistent with the observed correlations between the relative weight of inputs and other plant-level outcomes. The estimates of the elasticity of substitution are less than one, suggesting complementarities between capital and labor. This complementarity is plausible considering that the industries in the SMT are more likely to encapsulate higher levels of automation. The estimates also imply that the production labor share is decreasing and the capital-production labor ratio is increasing in the relative price of production labor. In addition, the estimated total factor productivity distribution implied by CES production with endogenous technology choice differs significantly from standard models. The findings also suggest that the form of the production function matters for the assessment of the relationship between productivity, labor share, and automation.

Our simple exploration of dynamic relationships indicates that more automated plants tend to experience larger five- and ten-year declines in labor share and bigger increases in labor productivity. Given the positive connection between firm productivity and growth established in previous studies, these

findings are consistent with a model where the adoption and use of automation is behind the success of large and productive firms with low labor shares. Future work should consider this hypothesis in more detail.

References

Acemoglu, Daron and Pascual Restrepo. Robots and Jobs: Evidence from US Labor Markets. Working Paper 23285, National Bureau of Economic Research, 2017.

Acemoglu, Daron and Pascual Restrepo. Modeling Automation. *American Economic Review Papers and Proceedings*, 108:48–53, May 2018a.

Acemoglu, Daron and Pascual Restrepo. Artificial Intelligence, Automation, and Work. Working paper, Federal Trade Commission, 2018b.

Arellano, Manuel and Stephen Bond. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2):277–297, April 1991.

Autor, David and Anna Salomons. Is Automation Labor-Displacing? Productivity Growth, Employment, and the Labor Share. *Brookings Papers on Economic Activity*, (Spring):1–87, 2018.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. The Fall of the Labor Share and the Rise of Superstar Firms. Working Paper 23396, National Bureau of Economic Research, May 2017a.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. Concentrating on the Fall of the Labor Share. *American Economic Review*, 107(5):180–85, May 2017b.

Autor, David H., David Dorn, and Gordon H. Hanson. The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, 103(6):2121–68, October 2013.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. R&D investment, exporting, and productivity dynamics. *American Economic Review*, 101(4):1312–44, June 2011.

Barkai, Simcha. Declining Labor and Capital Shares. Working paper, University of Chicago, 2016.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. Cross-Country Differences in Productivity: The Role of Allocation and Selection. *American Economic Review*, 103(1):305–34, February 2013.

Beede, David and Kan H. Young. Patterns of advanced technology adoption and manufacturing performance: Employment growth, labor productivity, and employee earnings. 05 1996.

Bloom, Nicholas. The impact of uncertainty shocks. *Econometrica*, 77(3): 623–685, 2009.

Caballero, Ricardo J., Eduardo M. R. A. Engel, and John C. Haltiwanger. Aggregate Employment Dynamics: Building from Microeconomic Evidence. *American Economic Review*, 87(1):115–37, March 1997.

Cooper, Russell W. and John C. Haltiwanger. On the Nature of Capital Adjustment Costs. *Review of Economic Studies*, 73(3):611–633, 2006.

De Loecker, Jan. Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, 79(5): 1407–1451, 09 2011.

Dinlersoz, Emin, Jeremy Greenwood, and Henry Hyatt. What Businesses Attract Unions? Unionization over the Life Cycle of U.S. Establishments. *Industrial and Labor Relations Review*, 70(3):733–766, 2017.

Doms, Mark, Timothy Dunne, and Mark J. Roberts. The Role of Technology Use in the Survival and Growth of Manufacturing Plants. *International Journal of Industrial Organization*, 13:523–542, 12 1995.

Doms, Mark, Timothy Dunne, and Ken Troske. Workers, Wages, and Technology. *Quarterly Journal of Economics*, 112(1):253–290, February 1997.

Doraszelski, Ulrich and Jordi Jaumandreu. Measuring the bias of technological change. *Journal of Political Economy*, 126(3):1027–1084, 2018.

Dunne, Timothy. Plant Age and Technology Use in U.S. Manufacturing Industries. *RAND Journal of Economics*, 25(3):488–499, Autumn 1994.

Dunne, Timothy and James A. Schmitz Jr. Wages, Employment Structure and Employer Size−Wage Premia: Their Relationship to Advanced-Technology Usage at US Manufacturing Establishments. *Economica*, 62:89–107, 02 1995.

Elsby, Michael, Bart Hobijn, and Aysegul Sahin. The Decline of the U.S. Labor Share. *Brookings Papers on Economic Activity*, 44(2 (Fall)):1–63, 2013.

Ericson, Richard and Ariel Pakes. Markov-Perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62(210):53–82, 1995.

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. Firm-Level Dispersion in Productivity: Is the Devil in the Details? *American Economic Review*, 106(5):95–98, May 2016.

Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. Macro and Micro Dyamics of Productivity: Is the Devil in the Details? Working Paper 23666, National Bureau of Economic Research, August 2017.

Gandhi, Amit, Salvador Navarro, and David Rivers. On the identification of production functions: How heterogeneous is productivity? Working paper, 2016.

Graetz, Georg and Guy Michaels. Robots at work. *The Review of Economics and Statistics*, 100(5):753–768, 2018.

Haltiwanger, John and Zoltan Wolf. Declining Reallocation in the U.S.: Implication for Productivity Growth? Unpublished manuscript, 2018.

Karabarbounis, Loukas and Brent Newman. The Global Decline of the Labor Share. *The Quarterly Journal of Economics*, 129(1):61–103, 2014.

Kelley, Maryellen R. and Todd A. Watkins. Are defense and non-defense manufacturing practices all that different? *The Defense Industry in the Post-Cold War Era: Corporate Strategies and Public Policy Perspectives*, 1998.

Lawrence, Robert. Recent Declines in Labor's Share in US Income: A Preliminary Neoclassical Account. Working Paper 21296, National Bureau of Economic Research, 2015.

Levinsohn, James A. and Amil Petrin. Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70 (2):317–341, April 2003.

McGuckin, Robert, Mary Streitweiser, and Mark Doms. Advanced Technology Usage and Productivity Growth. *Economics of Innovation and New Technology*, 7(1):1–26, 1998.

Oberfield, Ezra and Devesh Raval. Micro Data and Macro Technology. Working paper, Federal Trade Commission, 2014.

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

Raval, Devesh. The Micro Elasticity of Substitution and Non-neutral Technology. Working paper, Federal Trade Commission, 2017.

Table 1: Descriptive statistics of SMT industries and technologies

(a) Summary 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
Avg. (SMT industries)		0.111	0.106	0.062	0.568	0.989
Avg. (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 manufacturing industries outside the five SMT industries.

(b) Diffusion rates of technologies covered in the SMT

	Diffusion Rate (%)	
Technology	1988	1993
Group 1. Fabrication and Machining		
Numerically-controlled/computer-numerically-controlled (NC/CNC) Machines	41.4	46.9
Flexible Manufacturing Cells or Systems	10.7	12.7
Materials Working Laser	4.3	5.0
Pick and Place Robot	7.7	8.6
Other Robot	5.7	4.8
Group 2. Design and Engineering		
Computer-Aided Design/Engineering	39.0	58.8
Computer-Aided Manufacturing	16.9	25.6
Digital Data Representation	9.9	11.3
Group 3. Inspection and Quality Control		
Computers used for Control on the Factory Floor	27.3	26.9
Factory Network	16.2	22.1
Programmable Controller	32.1	30.4
Technical Data Network	18.9	29.3
Intercompany Network Linking Plant to Suppliers/Customers/Subcontractors	14.8	
Automated Sensor-Based Inspection/Testing:		
Incoming or In-Process Materials	10.0	9.9
Final Product	12.5	12.5
Group 4. Materials Handling		
Automatic Guided Vehicle System	1.5	1.1
Automatic Storage and Retrieval System	3.2	2.6

Source: Survey of Manufacturing Technology printed summaries from Current Industrial Reports SMT(88)-1 and SMT(93)-3

Table 2: Questions on technology use and investment in 1991 Survey of Manufacturing Technology

Survey question number and text	Survey Response	Recoded Response
1. What degree do the manufacturing operations in this plant depend on technologically advanced equipment and software?	Not applicable	0
	$< 10\%$	1
	10% to 25%	2
	25% to 49%	3
	50% to 74%	4
	$\geq 75\%$	5
2. Indicate the range that best reflects this plant's total investment in technologically advanced equipment and software for the past three years. Exclude education and training but include plant modifications, construction, integration, and equipment and software purchased and developed.	Not applicable	0
	< \$100K	1
	\$100K-1M	2
	\$1M-5M	3
	\$5M-\$10M	4
	$\geq$ \$10M	5
11. What percentage of this plant's operations will depend upon technologically advanced equipment and software in three years?	Not applicable	0
	$< 10\%$	1
	10% to 25%	2
	25% to 49%	3
	50% to 74%	4
	$\geq 75\%$	5
12. What are your plans to acquire technologically advanced equipment and software for this plant over the next three years?	Not applicable	0
	Under consideration	1
	Minor upgrade ($< 25\%$)	2
	Major upgrade (25%-75%)	3
	Total replacement ($\geq 75\%$)	4

Source: Survey of Manufacturing Technology survey form - Current Industrial Reports SMT(91)-2, Appendix A

Table 3: Technology and plant outcomes

(a) Labor share and plant characteristics

	Labor Share			Fraction of Workers in Production
	All	Production	Non-production	
technology index I	-0.023***	-0.041***	0.026	-0.035***
	[0.008]	[0.016]	[0.020]	[0.009]
R^2	0.27	0.28	0.31	0.30
technology index II	-0.052***	-0.083***	-0.0064	-0.041***
	[0.014]	[0.019]	[0.022]	[0.011]
R^2	0.27	0.28	0.31	0.30
N	8100	8100	8100	8100

(b) Labor productivity and plant characteristics

	Labor Productivity		
	All	Production	Non-production
technology index I	0.090***	0.120***	0.021
	[0.018]	[0.021]	[0.026]
R^2	0.26	0.30	0.23
technology index II	0.120***	0.142***	0.055
	[0.026]	[0.029]	[0.034]
R^2	0.26	0.30	0.23
N	8100	8100	8100

(c) Salaries and wages per employee and plant characteristics

	Average Wage		
	All	Production	Non-production
technology index I	0.073***	0.084***	0.044***
	[0.013]	[0.012]	[0.013]
R^2	0.22	0.24	0.09
technology index II	0.083***	0.086***	0.048***
	[0.015]	[0.014]	[0.016]
R^2	0.22	0.24	0.08
N	8100	8100	8100

Notes: All continuous variables in logs. Standard errors –in parentheses– 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 2. Technology index II is based only on the investment question (Question 2). All regressions include other plant characteristics as controls: five plant size (employment) categories (1-20 emp, 20-99 emp, 100-499 emp, 500-999 emp, 1000+ emp), four plant age categories (0-5 yrs, 5-14 yrs, 15-29 yrs, 30+ yrs), a production worker unionization indicator (1 if the plant has a union contract for production workers), export intensity indicator (1 if more than 50% of the plant's products are exported), an indicator of military production (1 if the plant is engaged in production to military specs), a foreign-ownership indicator (1 if 10% or more of the voting stock or other equity rights are foreign-owned), an indicator of shipment to defense agencies (1 if the plant ships directly to DOD or Armed Services), an indicator of shipment to primary contractors for defense agencies (1 if shipments are made to a primary defense contractor), and 4-digit SIC industry fixed effects. N is rounded for disclosure avoidance.

Table 4: Technology and the growth in plant outcomes

(a) Dependent variable: growth in production labor share

	1997	2002	1997	2002
technology index I	-0.080***	-0.075***	–	–
	[0.014]	[0.019]		
technology index II	–	–	-0.078***	-0.067***
			[0.013]	[0.020]
employment growth 1997	0.133***	–	0.106***	–
	[0.020]		[0.019]	
employment growth 2002		0.171***		0.170***
		[0.018]		[0.018]
R^2	0.02	0.04	0.05	0.04
N	6400	5200	6400	5200

(b) Dependent variable: growth in production labor productivity

	1997	2002	1997	2002
technology index I	0.097***	0.072***		
	[0.013]	[0.018]		
technology index II			0.106***	0.067***
			[0.013]	[0.019]
employment growth 1997	-0.168***		-0.164***	
	[0.019]		[0.019]	
employment growth 2002		-0.158***		-0.157***
		[0.018]		[0.018]
R^2	0.04	0.04	0.04	0.04
N	6400	5200	6400	5200

Notes: See Notes to Table 3.

Table 5: Elasticity estimates

(a) Estimates of σ

	SMT (IV)		ASM ($\text{GMM}_{t-2,t-p}$)	
	full	simple	$p = 7$	$p = 8$
$\widehat{\sigma}$	0.63***	0.71***	0.60***	0.38***
$\widehat{\rho}=\frac{\widehat{\sigma}-1}{\widehat{\sigma}}$	-0.59	-0.41	-0.67	-1.63
N	4400	4400	11500	5500

Notes: IV: cross-section IV with and without capital in the regression. $\text{GMM}_{t-2,t-p}$: GMM using indicated lagged differences as instruments. These regressions are based on the earliest possible lags available where the Hansen test of overidentifying restrictions do not reject the null of orthogonality. N is rounded for disclosure avoidance.

(b) Production function estimates

	N	$\widehat{\gamma}$	β_1	β_2	β_3	$\sum_j \widehat{\beta}_j + \widehat{\gamma}$
			FOC			
	4100	0.22	0.12	0.43	0.02	0.77
		(0.002)	(0.002)	(0.003)	(0.001)	(0.004)
			Equation (15)			
NLS, α_i	4000	0.17	0.12	0.43	0.02	0.73
		(0.075)	(0.002)	(0.003)	(0.001)	(0.074)
NLS, $\alpha_i = 1/2$	4000	0.25	0.12	0.43	0.02	0.81
		(0.047)	(0.002)	(0.003)	(0.001)	(0.047)

Notes: Standard errors are bootstrapped. All elasticities are based on output and input distributions from which outliers are removed. Variable input elasticities are fixed across specifications. N is rounded for disclosure avoidance.

Table 6: Properties of implied productivity distributions

(a) Descriptive statistics, 1991

	N	stdev	skewness	kurtosis
CD_{CRS}	4000	0.43	-0.74	9.71
CD_{NCRS}	4000	0.48	-0.60	6.31
CES_{FOC}	4000	0.59	-0.87	5.64
CES_{EN}	4000	0.58	-0.73	5.73
CES_{EX}	4000	0.59	-0.97	5.87

(b) Correlations

	CD_{CRS}	CR_{NCRS}	CES_{FOC}	CES_{EN}	CES_{EX}
			1991		
CD_{CRS}	1				
CD_{NCRS}	0.86	1			
CES_{FOC}	0.82	0.91	1		
CES_{EN}	0.80	0.94	0.99	1	
CES_{EX}	0.82	0.88	0.99	0.97	1
			1992		
CD_{CRS}	1				
CD_{NCRS}	0.82	1			
CES_{FOC}	0.78	0.9	1		
CES_{EN}	0.77	0.94	0.99	1	
CES_{EX}	0.79	0.87	0.99	0.98	1

(c) P-values from the Kolmogorov-Smirnov test

x	y	H_0: x=y	H_0: x<y	H_0: x>y
CD_{CRS}	CD_{NCRS}	0	0	0
CD_{CRS}	CES_{FOC}	0	0	0
CD_{CRS}	CES_{EN}	0	0	0
CD_{CRS}	CES_{EX}	0	0	0
CD_{NCRS}	CES_{FOC}	0	0	0
CD_{NCRS}	CES_{EN}	0	0	0
CD_{NCRS}	CES_{EX}	0	0	0
CES_{EN}	CES_{EX}	0.04	0.02	0.13
CES_{FOC}	CES_{EN}	0.83	0.75	0.46

Notes: Outliers are filtered in yearly distributions. Industry-year effects are removed. N is rounded for disclosure avoidance. Based on the K-S test, we reject all three H_0s for any pair of CD and CES residuals, irrespective of how the residuals were calculated.

Table 7: The relationship between automation, productivity and labor share

(a) Production Labor Share of Revenue

	I	II	III	IV	V
CES_{EN}	-0.177***	-0.337***	-0.353***	-0.335***	-0.350***
	[0.022]	[0.029]	[0.029]	[0.029]	[0.029]
technology index I		-0.042**	-0.034**		
		[0.021]	[0.017]		
technology index II				-0.080***	-0.078***
				[0.028]	[0.027]
employment		0.115***	0.099***	0.125***	0.109***
		[0.013]		[0.013]	[0.013]
R^2	0.02	0.04	0.09	0.05	0.09
N	4600	4600	4600	4600	4600

(b) Production labor share of composite input expenditures

	I	II	III	IV	V
CES_{EN}	0.184***	0.299***	0.298***	0.303***	0.302***
	[0.013]	[0.018]	[0.018]	[0.018]	[0.018]
technology index I		-0.072***	-0.071***		
		[0.014]	[0.014]		
technology index II				-0.127***	-0.126***
				[0.015]	[0.015]
employment		-0.063***	-0.064***	-0.049***	-0.051***
		[0.008]	[0.008]	[0.008]	[0.008]
R^2	0.07	0.12	0.13	0.13	0.14
N	4600	4600	4600	4600	4600

(c) Technology index

	Technology Index I			Technology Index II		
	I	II	III	I	II	III
CES_{EN}	0.279***	0.028**	0.029**	0.361***	0.049***	0.048***
	[0.013]	[0.014]	[0.014]	[0.013]	[0.014]	[0.014]
employment		0.168***	0.171***		0.208***	0.208***
		[0.007]	[0.007]		[0.006]	[0.006]
R^2	0.10	0.22	0.23	0.17	0.36	0.36
N	4600	4600	4600	4600	4600	4600

Notes: See Notes to table 3. Specifications III and V in panels (a) and (b) include plant characteristics described in Table 3. Productivity, technology indices and employment are expressed as deviations from industry means. N is rounded for disclosure avoidance.

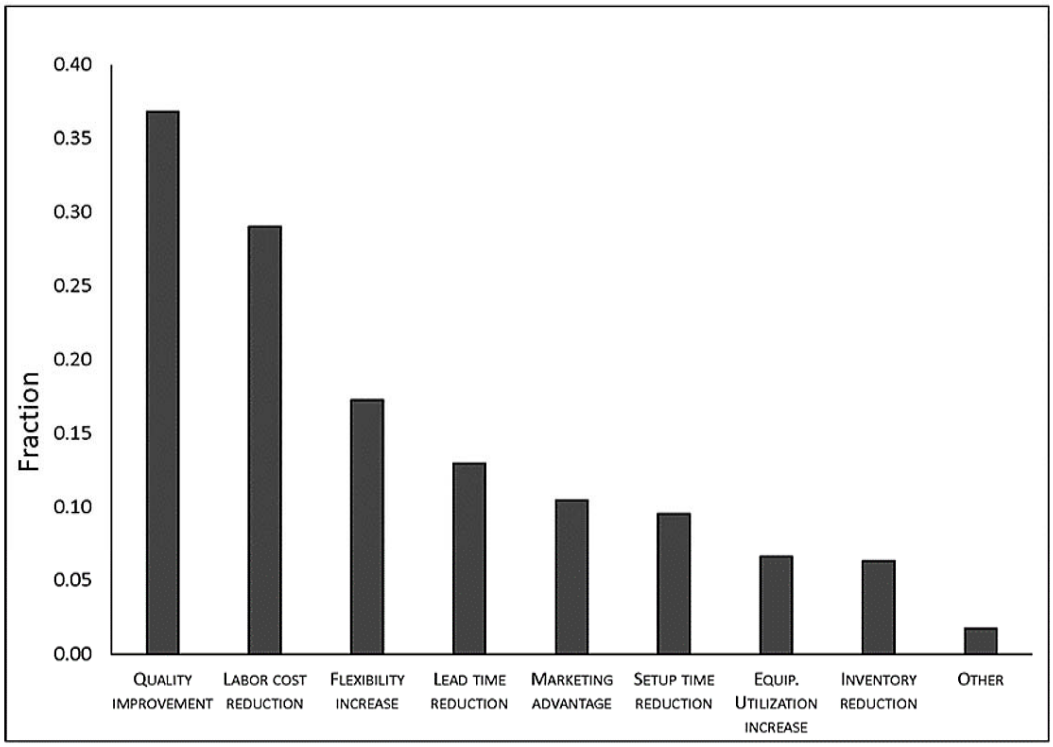
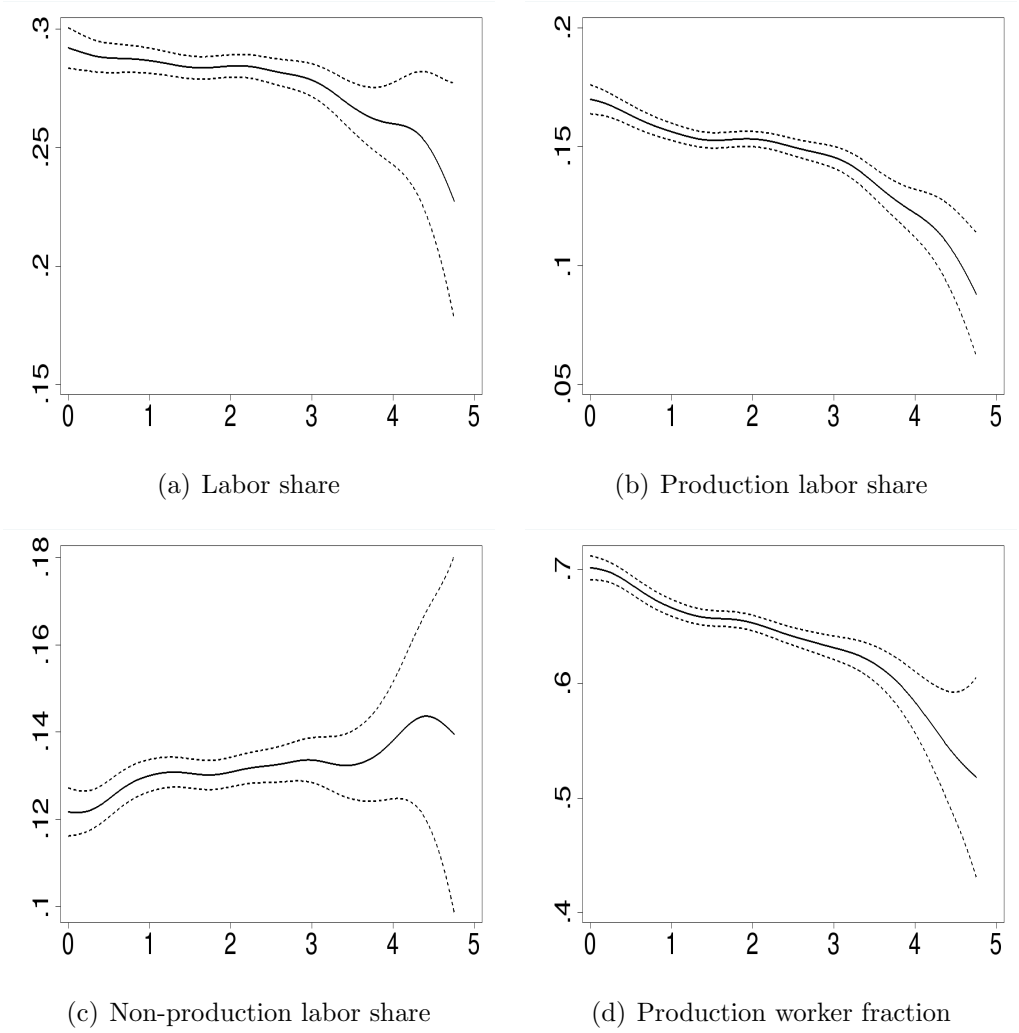


Figure 1: Benefits derived from the use of automation-related technologies, subjective assessment of plants.

Source: Survey of Manufacturing Technology



(a) Labor share

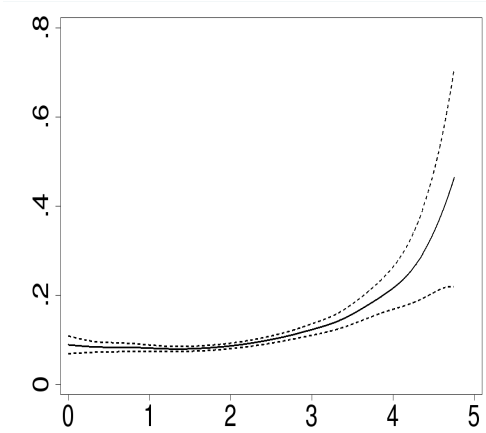
(b) Production labor share

(c) Non-production labor share

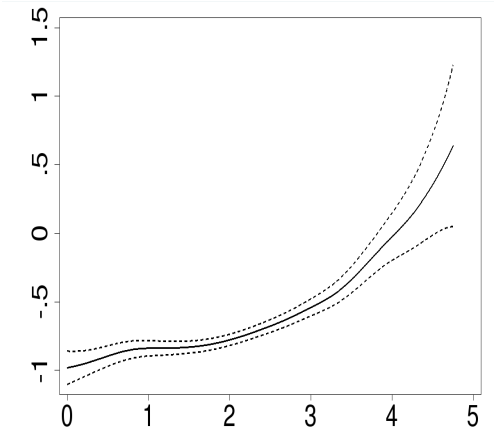
(d) Production worker fraction

Figure 2: Labor usage as a function of technology index.

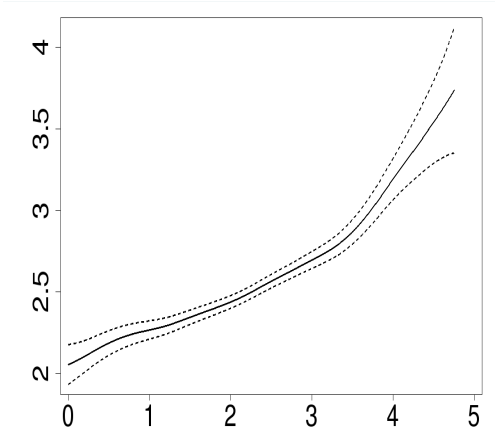
Labor share is the share of labor costs in value of shipments. The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Dotted lines show 95% confidence intervals for local polynomial smoothing. Source: Survey of Manufacturing Technology, Annual Survey of Manufactures, Census of Manufacturing



(a) Capital share



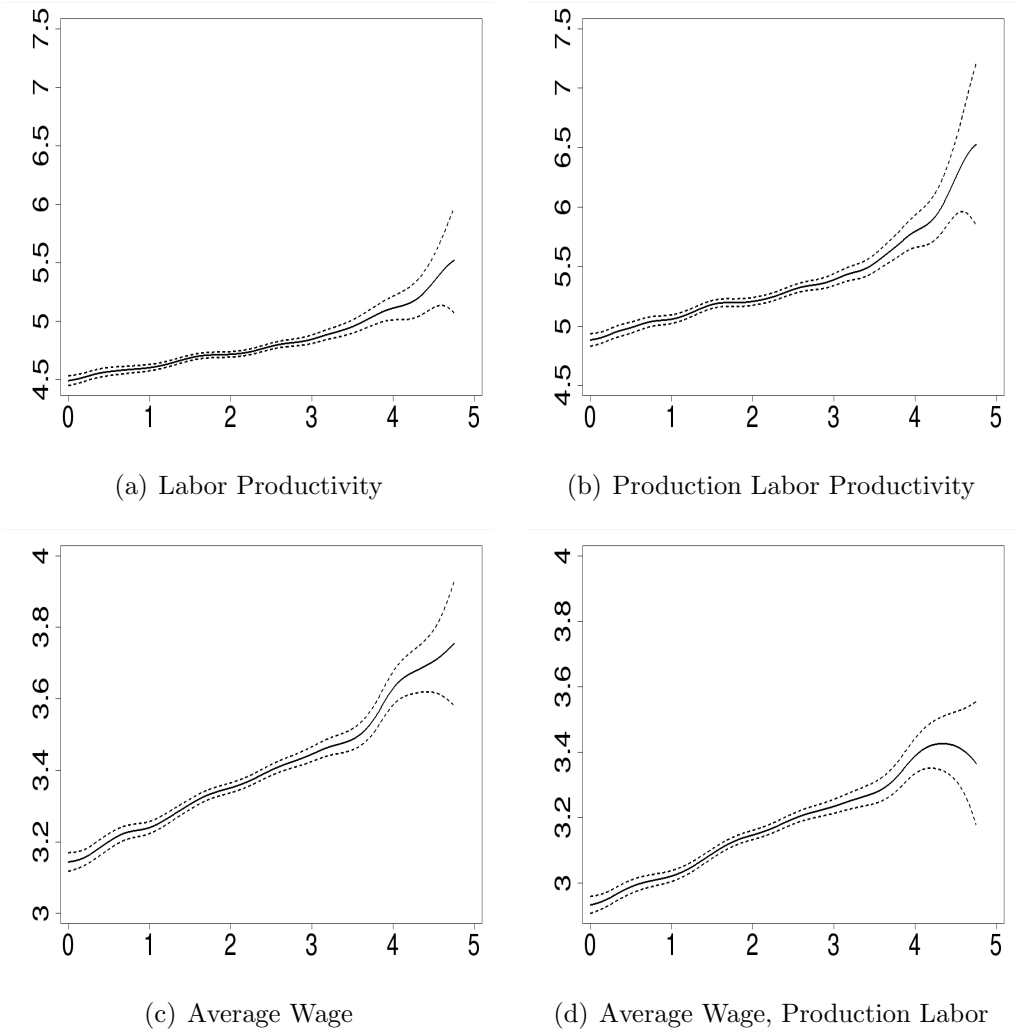
(b) Capital-production labor cost ratio, logs



(c) Capital-production labor ratio, logs

Figure 3: Capital usage as a function of technology index.

Capital share is the share of capital costs in value of shipments. The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Dotted lines show 95% confidence intervals for local polynomial smoothing. Source: Survey of Manufacturing Technology, Annual Survey of Manufactures, Census of Manufacturing



(a) Labor Productivity

(b) Production Labor Productivity

(c) Average Wage

(d) Average Wage, Production Labor

Figure 4: Value of shipments per worker (in logs) and average wage as a function of technology index.

The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Revenue is measured as Total Value of Shipments. Dotted lines show 95% confidence intervals for local polynomial smoothing. Source: Survey of Manufacturing Technology, Annual Survey of Manufactures, Census of Manufacturing

A Appendix

A.1 Evolution of SMT Industries

In order to shed more light on the general evolution of some key indicators for the SMT industries, Figures A2(a)-A2(c) show industry-level factor share measures for the five industries over the period 1972-1997.[58] Both overall and production labor shares decline in the 5 industries during this period, see Figures A2(a)-A2(b). The capital share is relatively stable across all industries, with some increase observed in SIC 34 and 36, see Figure A2(c). These dynamics are reflected in the ratio of capital share to production labor share, see Figure A2(d). The 5-factor TFP measures in Figure A3, suggest that productivity is higher and grows faster in more capital intensive industries (notably, SIC 35 and 36). Overall, the trends are in line with the general decline in labor share in manufacturing, and the five industries tend to exhibit largely similar trends.

A.2 Optimization

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

[58] We use industry-level data from the NBER-CES Manufacturing Industry Database, available at http://www.nber.org/nberces/. This data was supplemented with industry-level capital prices from the Bureau of Labor Statistics.

A.3 Details of Production Function Estimation

Below is an outline of the estimation approach.

1. Use (11) to obtain an estimate of σ.
2. Estimate input elasticities using the average revenue share of the input $\widehat{\beta}_j = N^{-1} \sum_i \frac{w_{ji} X_{ji}}{R_i}$.[59]
3. Net out the contribution of variable input costs from revenue: $\widehat{\mathcal{B}}_i = r_i - \sum_j \widehat{\beta}_{ji} w_{ji} X_{ji}$.
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
$$\ln \widehat{T}_i \equiv \frac{1}{\widehat{\rho}} \ln \left[\alpha_i^{2/\widehat{\sigma}} K_i^{\widehat{\rho}} + (1-\alpha_i)^{2/\widehat{\sigma}} L_{pi}^{\widehat{\rho}} \right]. \tag{13}$$
5. Determine the joint contribution of state variables and the proxy by estimating
$$\widehat{\mathcal{B}}_i = \phi(Z_i, p) + v_i, \tag{14}$$
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 $Z_i = \left(\ln \widehat{T}_i, \bar{t}_i \right)'$.[60]
6. Given fitted values $\widehat{\phi}_{it}$ from equation (14), use nonlinear least squares to estimate
$$\widehat{\mathcal{B}}_{it} = \delta_T \ln \widehat{T}_{it} + h\left(\widehat{\phi}_{it-1} - \delta_T \ln \widehat{T}_{it-1} \right) + \nu_{it}. \tag{15}$$

[59] Averaging mitigates the effects of measurement error, and is often used in the empirical productivity literature.

[60] Treating T_i as a state variable is 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.

where h is a second-order polynomial in its argument. Under the assumptions of the model, δ_T in regression (15) 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 with $\ln \widehat{T}_i$, and plant age they determine $\widehat{\phi}_i$ in (14). Under the assumptions of the model, this value controls for unobserved productivity differences across plants when estimating δ_T using data from 1992 in (15). If the plant-level productivity process is Markovian –a standard assumption in the empirical productivity literature– then δ_T is consistently estimated in regression (15). The standard error of $\widehat{\delta}_T$ is obtained using bootstrap.

A.4 Specification Error

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

$$Q_i = \theta_i K_i^{\beta} L_{pi}^{1-\beta}, \tag{16}$$

when the underlying data generating process is a CES production function with decreasing returns to scale ($\gamma < 1$) and endogenous technology choice

$$Q_i = \theta_i [\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}]^{\gamma/\rho}. \tag{17}$$

For ease of exposition, the two production functions above abstract from the variable inputs L_n, M and E used in (2). Including them does not change the main conclusions regarding the theoretical relationship between specification error and technology. In addition, if the sampling error is small, their contribution to the specification error is negligible.[61]

[61] However, including these three inputs matter for the estimated specification error if their elasticities are estimated based on different methods for the CES production function used in the main text and for the CD specification.

Let $k_i = \ln K_i$, $l_{pi} = \ln L_{pi}$, and $\widehat{\sigma} = 1/(1-\widehat{\rho})$. The difference between the estimated CD-based log TFP and the estimated CES-based log TFP is given by

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

After some manipulation of terms, one can rewrite Δ_i as a sum of three components

$$\begin{aligned}
\Delta_i &= \left(\frac{\widehat{\gamma}}{\widehat{\rho}} \ln[\alpha_i^{2/\widehat{\sigma}} K_i^{\widehat{\rho}} + (1-\alpha_i)^{2/\widehat{\sigma}} L_{pi}^{\widehat{\rho}}] - \frac{\gamma}{\rho} \ln[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}] \right) \\
&\quad + \left(\frac{\gamma}{\rho} \ln[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}] - \beta k - (1-\beta) l_{pi} \right) \\
&\quad + (\widehat{\beta} - \beta)(k_i - l_{pi}) \\
&= \Delta_{i,CES}^E + \Delta_i^S - \Delta_{i,CD}^E,
\end{aligned}$$

where Δ_i^S is the specification error due to functional form assumption, and $\Delta_{i,CES}^E$ and $\Delta_{i,CD}^E$ are the estimation (sampling) errors associated with the CES and CD specifications, respectively.

Consider the specification error

$$\Delta_i^S = \frac{\gamma}{\rho} \ln[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_p^{\rho}] - \beta k_i - (1-\beta) l_{pi}. \tag{18}$$

If the estimation errors $\Delta_{i,CES}^E$ and $\Delta_{i,CD}^E$ are small, Δ_i is closely approximated by an estimate of Δ_i^S

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

In (18), a first order Taylor series approximation to the term $\ln[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_p^{\rho}]$ around $\rho = 0$ yields

$$\ln[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}] = \{B_i k + (1-B_i) l_{pi} - 2\left[B_i \ln \alpha_i + (1-B_i) \ln(1-\alpha_i)\right]\} \rho + \xi,$$

where

$$B_i = \frac{\alpha_i^2}{\alpha_i^2 + (1-\alpha_i)^2}, \tag{19}$$

and ξ is the approximation error. One can thus approximate the specification error, Δ_i^S, as follows

$$\widetilde{\Delta_i^S} = (\gamma B_i - \beta)(k_i - l_{pi}) + (\gamma - 1) l_{pi} - 2\gamma \left[B_i \ln \alpha_i + (1 - B_i) \ln (1 - \alpha_i) \right]. \tag{20}$$

The first two terms in the final expression indicate that the magnitude of the error depends on the capital-production labor ratio and production labor. To study the contribution of the third term, define

$$D(\alpha_i) = -2 \left[B_i \ln \alpha_i + (1 - B_i) \ln (1 - \alpha_i) \right]$$

The third term is then $\gamma D(\alpha_i)$. $D(\alpha_i)$ is always positive (since $\alpha_i \in (0,1)$), and it is a non-monotonic function of α_i. It achieves its maximum value of 1.386γ at $\alpha_i = 0.5$, and its minimum value of zero at $\alpha_i = 0$ or $\alpha_i = 1$. The shape of $D(\alpha_i)$ is shown in figure A1.

The term $\gamma D(\alpha_i)$ is bounded from above by 1.386γ, and its contribution to the specification error will be dominated by the first two terms, when γ is small.[62] The rest of the appendix hence assumes away the contribution of $\gamma D(\alpha_i)$.

What is the sign of the specification error, Δ_i^S? Consider the approximation, $\widetilde{\Delta_i^S}$, in (20). Because $\gamma < 1$, the second term in (20) is negative (for $l_{pi} > 0$, or equivalently, $L_p > 1$). For $(k_i - l_{pi}) > 0$ (i.e. $(K_i/L_{pi}) > 1$), the first term in (20) can be positive or negative, depending on whether $\gamma B_i - \beta$ is positive or negative.[63] If the first term is negative, $\widetilde{\Delta_i^S}$ is negative for all plants with at least one production worker and capital-production labor ratio greater than one.[64] In practice, the estimated specification error turns out to

[62] Note that the estimated value of γ is in the range $(0.17, 0.25)$ based on the samples used in this paper. These values imply a range of $(0, 0.34)$ for $\gamma D(\alpha_i)$.

[63] Because $B_i \in (0,1)$, one sufficient condition for $\gamma B_i - \beta$ to be negative for all $B_i \in (0,1)$ is $\gamma < \beta$.

[64] This ratio exceeds one in the samples used in this study.

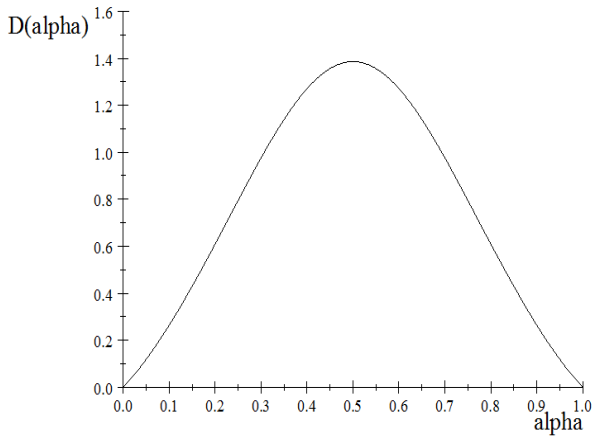


Figure A1: The shape of $D(\alpha_i)$.

be negative for nearly the entire set of plants.

Another question is whether the specification error is exacerbated for plants with higher degree of automation. Consider the first term in $\widetilde{\Delta_i^S}$. Both $(\gamma B_i - \beta)$ and $(k_i - l_{pi})$ are increasing functions of α_i when $\sigma < 1$, by the definition of B_i in (19), and by equations (3) and (4). Now, let $F(\alpha_i) = \gamma B_i - \beta$ and $G(\alpha_i) = k_i - l_{pi}$. Note that $F' > 0$ and $G' > 0$ given the preceding discussion. Then, the rate of change of first term with α_i is given by $F'G + G'F$. When $F > 0$, $F'G + G'F > 0$. When $F < 0$, $F'G + G'F$ can be positive or negative. As a result, the first term $(\gamma B_i - \beta)(k_i - l_{pi})$ can be an increasing or decreasing function of α_i. The second term in (20), $(\gamma - 1)l_{pi}$, can also be increasing or decreasing in α_i, depending on how l_{pi} changes with α_i. The model has no prediction on the direction of this change. For instance, $(\gamma - 1)l_{pi}$ is decreasing in α_i if large plants (large l_{pi}) are also the ones with higher α_i.[65] The overall sign of the change in the specification error as α_i increases then depends on

[65] This connection finds support in the SMT sample.

the behavior of the first and second terms.

Table A1: The relationship between Δ 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/prod. labor			0.064***	0.113***		
			[0.008]	[0.004]		
technology index I					0.056***	0.055***
× capital/prod. 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: See notes to Table 3. Specification VI includes other plant characteristics aside from employment.

Empirical results indicate that the specification error becomes more negative as automation increases. In other words, specification error tends to be larger (in absolute value) for more automated plants, and the CD_{CRS} tends to underestimate the underlying TFP (as estimated by CES_{EN}) more for such plants. Table A1 contains the projections of $\Delta_i = \text{CD}_{\text{CRS}}\text{-CES}_{\text{EN}}$ on key components of Δ_i: technology index (a proxy for α_i), production labor (l_{pi}), capital-production labor ratio ($k_i - l_{pi}$), and an interaction of the technology index with the capital-production labor ratio, all expressed as deviations from 4-digit SIC industry means. The interaction of the technology index with the capital-production labor ratio is a proxy for the term $(\gamma B_i - \beta)(k_i - l_{pi})$ in expression (20). The coefficient estimates for the bivariate projections in Table A1 indicates that Δ_i decreases as the technology index or production labor increases, but increases as capital-production labor ratio increases (Specifications I-III). These relationships also hold when all three variables are used together in the projection (Specification IV). In addition, controlling for production labor, Δ_i is positively associated with the interaction of the technology index and the capital-production labor ratio (Specifications V and VI).

A.5 Additional Results

Table A2: The relationship between 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 Lamda	-0.005	-0.078*	-0.008	-0.074*
N	8100	8100	8100	8100

Notes: All continuous variables in logs. Standard errors are clustered by 4-digit SIC industry. (*), (**), (***) indicate significance at 10, 5, and 1% level, respectively. The coefficient estimates are based on the Heckman two-step correction. Technology index I is based on all 4 survey questions in Table 2. 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

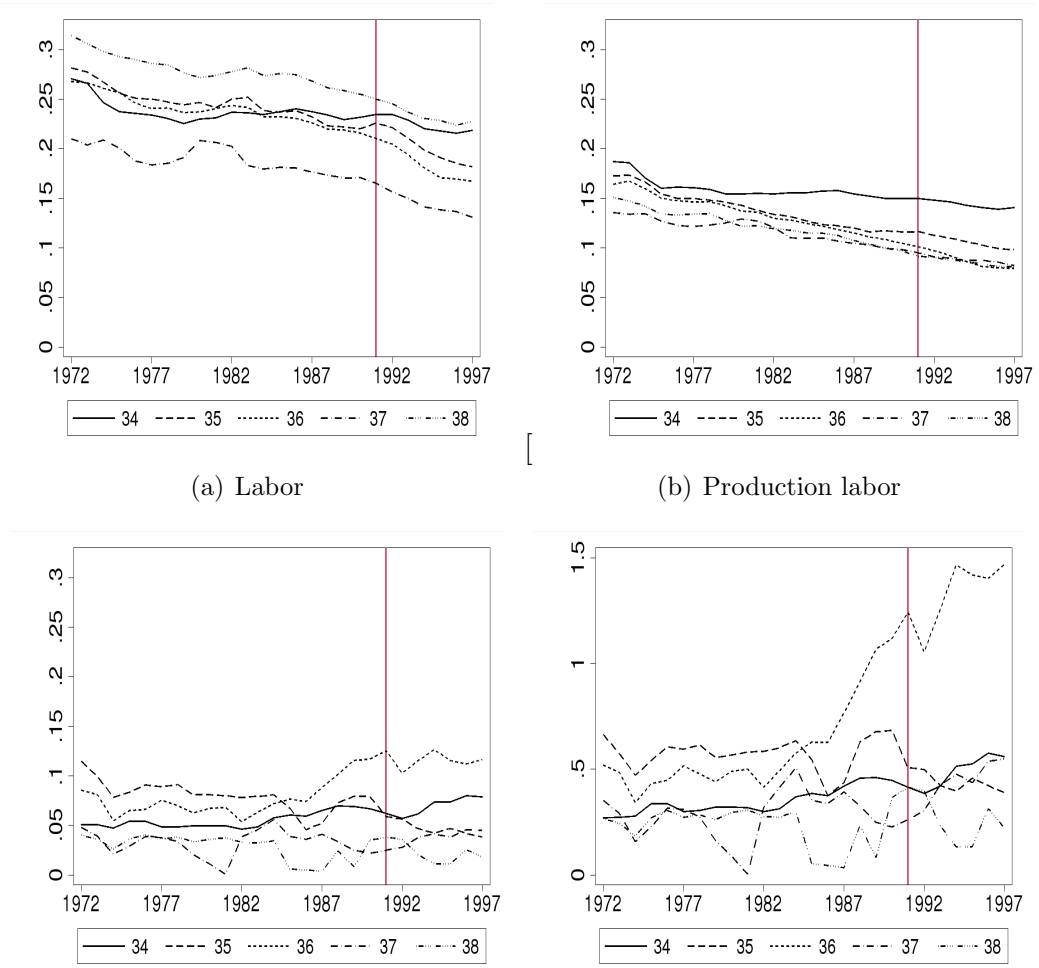
	Growth in Production Labor Productivity			
	1997	2002	1997	2002
technology index I	0.096***	0.090***		
	[0.014]	[0.019]		
technology index II			0.105***	0.086***
			[0.014]	[0.020]
employment growth 1997	-0.168***		-0.168***	
	[0.012]		[0.012]	
employment growth 2002		-0.157***		-0.156***
		[0.012]	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

Notes: All continuous variables in logs. 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 2. 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 A4: The relationships 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: All continuous variables in logs. Standard errors are clustered by 4-digit SIC industry. (*), (**), (***) indicate significance at 10, 5, and 1% level, respectively. The coefficient estimates are based on bivariate regressions. Technology index I is based on all 4 survey questions in Table 2. 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.



(a) Labor

(b) Production labor

(c) Capital

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

Figure A2: The evolution of the shares of capital and labor costs in the total value of shipments, in per cent.

The two-digit SIC codes denote the following industries: Fabricated Metal Products (34), Industrial Machinery and Equipment (35), Electronic and Other Electric Equipment (36), Transportation Equipment (37), and Instruments and Related Products (38). Vertical lines indicate survey year (1991). Source: NBER-CES database, 2-digit industries in the Survey of Manufacturing Technology. The capital stock measure is not quality-adjusted.

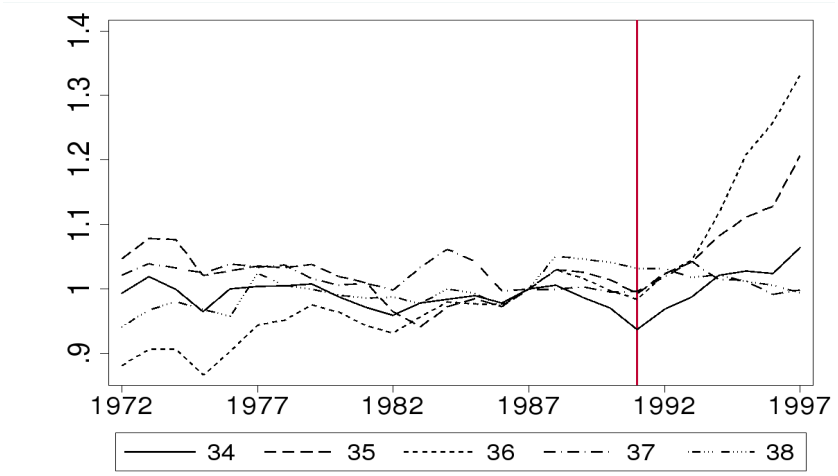
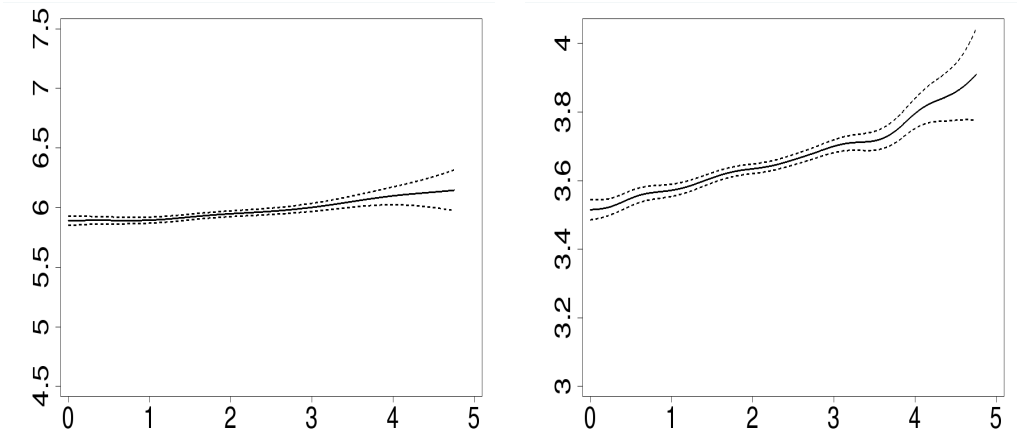


Figure A3: Average log-TFP across SIC industries

See notes to Figure A2.

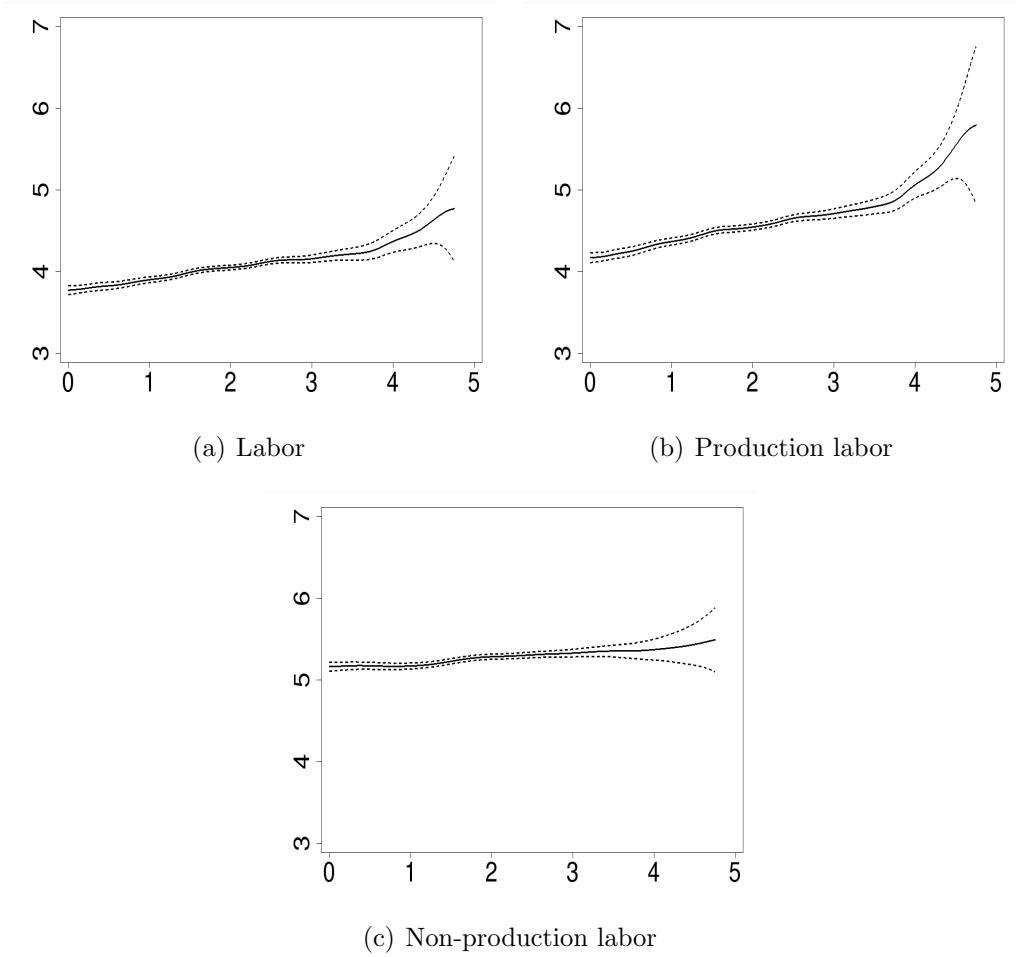


(a) Labor Productivity

(b) Average Wage

Figure A4: Non-production labor productivity and average wage (in logs) as a function of technology index.

The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Average wage is measured as payroll divided by the number of employees. Dotted lines show 95% confidence intervals of local polynomial smoothing. Source: Survey of Manufacturing Technology



(a) Labor

(b) Production labor

(c) Non-production labor

Figure A5: Value added per worker (in logs) as a function of technology index

The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Dotted lines show 95% confidence intervals for local polynomial smoothing. Source: Survey of Manufacturing Technology, Annual Survey of Manufactures, Census of Manufacturing