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

Emin Dinlersoz† — U.S. Census Bureau

Zoltan Wolf‡ — New Light Technologies

December 4, 2018
Preliminary

Abstract

This paper provides new evidence on the plant-level relationship between automation, labor and capital usage, and productivity. The evidence, based on the U.S. Census Bureau's Survey of Manufacturing Technology, indicates that more automated establishments have lower production labor share and higher capital share, and a smaller fraction of workers in production who receive higher wages. These establishments also have higher labor productivity and experience larger long-term labor share declines. The relationship between automation and relative factor usage is modelled using a CES production function with endogenous technology choice. This deviation from the standard Cobb-Douglas assumption is necessary if the within-industry differences in the capital-labor ratio are determined by relative input price differences. The CES-based total factor productivity estimates are significantly different from the ones derived under Cobb-Douglas production and positively related to automation. The results, taken together with earlier findings of the productivity literature, suggest that the adoption of automation may be one mechanism associated with the rise of superstar firms.

JEL Codes: D24, O33, J30, L60

Keywords: advanced manufacturing technology, automation, technology choice, total factor productivity, capital-labor substitution, labor share, CES production function, productivity estimation, 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 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 both the decline in employment, and the increase in output and productivity in U.S. manufacturing over the past decades, during which labor's share of output has also diminished. As robots and machines increasingly take over the tasks performed by humans, the reliance on labor can recede further. These aggregate trends notwithstanding, micro evidence on the connection between automation, labor share, and productivity, have been scarce, mainly due to a lack of detailed measures on the use and extent of automation at the plant level.

This paper provides new evidence on the nexus of automation, total factor productivity (TFP), and labor share using plant-level measures of automation from the U.S. Census Bureau's 1991 Survey of Manufacturing Technology (SMT). The SMT was designed to collect data on the adoption and use of automation-related advanced technologies, making it ideal for the type of analysis carried out here. The stylized facts, discussed in more detail in Section 2, point to a relationship between automation and relative factor usage that is consistent with theories emphasizing the potentially adverse effects of automation on labor engaged in production.[1] Specifically, overall labor share in the value of shipments is decreasing in the degree of automation mainly because the relationship between production labor share and automation is negative. In addition, more automated plants tend to have a lower fraction of their workers engaged in production and pay higher wages to production workers. Furthermore, plants with higher recent investment in automation experience larger declines in production labor share on a five-to-ten-year horizon. These patterns indicate a negative association between automation and the production labor share, both across plants and over time.

The stylized facts suggest that the differences in the capital-to-production-labor ratio and the ratio of expenditure shares for these two factors are non-trivial. More importantly, the variation in these measures is systematically related to the degree of automation, which points to a model of production that allows for within-industry variation in both the capital-labor ratio and relative factor shares.[2] Motivated by these observations, a general constant elasticity of substitution (CES) model of production is considered, in which the production unit adjusts the relative weights of capital and production labor in the input index as input prices vary. The variation in relative input price across plants is the main determinant that explains the differences in capital-labor ratio and the relative revenue shares of the two inputs. The sensitivity of relative usage of capital versus production labor to their prices is characterized by the elasticity of substitution between

[1] See, for instance, Acemoglu and Restrepo (2018a,c).

[2] The implications of assuming a Cobb-Douglas technology and competitive input markets are not consistent with these stylized facts. In particular, these assumptions imply that there is no relative input price variation across plants in the same industry, and consequently, the capital-labor ratio should be constant.

these two factors. Since this elasticity is a key technology parameter, the first step in the analysis is to estimate it.[3] Given an estimate of the elasticity, the remaining parameters of the production function are determined following a methodology similar to the one in Haltiwanger and Wolf (2018). The approach uses first-order conditions of the plant's optimization problem in order to determine the elasticity of variable factors. The elasticities of quasi-fixed factors are estimated controlling for unobserved TFP differences using plant-level variation in advanced technologies investment available from the 1991 SMT.

The elasticity of substitution estimates imply that the labor share declines as the price of production labor increases relative to the capital rental rate. Conditional on this estimate, the CES production function estimates yield a TFP distribution that is significantly different from the one implied by the standard Cobb-Douglas (CD) assumption with constant returns-to-scale, and other variants of the Cobb-Douglas and CES-based approaches. The findings also indicate that larger and more productive plants tend to rely more heavily on automation, and have lower production labor share. In other words, low production labor share is mainly a characteristic of larger, highly automated, and more productive plants.

This study is related to previous research on the role of relative factor-price differences that beget substitution away from production labor. Some of the prior studies use industry-level data and indirect inference to learn about the degree of diffusion of automation and its effects.[4] The empirical analysis in these studies relies on data on the relative price of equipment and the amount of certain types of capital, measured for broad industry aggregates.[5] In contrast, this paper uses direct micro-level measures of automation from the 1991 SMT. The objective of the SMT was to collect data from industries where the use of advanced technologies and automation is relatively more prevalent implying that the capital stock in these industries is more likely to be automation-related. This feature of the survey makes it ideal for studying patterns of capital-labor substitution.[6] The plant-level measures of automation used in this analysis cover four

[3]The parameter is identified using plant-level variation in labor usage, both in cross-section and over time. A similar methodology can be found in Raval (2017).

[4]See, e.g., Elsby et al. (2013), Karabarbounis and Newman (2014), and Graetz and Michaels (2015).

[5]For instance, Elsby et al. (2013) and Karabarbounis and Newman (2014) exploit the fall in the broad industry-level relative prices of capital to explain the decline in labor share. Acemoglu and Restrepo (2017) use data on the diffusion of robots available only by broad industry classifications to analyze local employment effects of automation.

[6]The majority of prior work in this literature utilizes general measures of capital stock, which arguably contain information on stocks of capital related to advanced technology. Previous research has also used measures of information technology investment and utilization reliance to study productivity (e.g., Brynjolfsson and Hitt; Brynjolsson and Hitt (2003), and Brynjolfsson and Yang (1996)). In its broader definition, automation is not limited to utilization of computers and IT, but also includes many types of robots and machines in which pre-programmed computer software dictates the movement of factory tools and machniery (CNC machines), metal-working lasers, optical inspection devices, automatic-guided vehicles, and many other technologies. Similar statements hold about labor. It is unlikely that all labor is substitutable with automation. Many types of labor, especially ranks of non-production labor such as managers, marketing and IT personnel may not face the same risk of displacement by automation as production labor. The findings in this paper are consistent with this view.

broad technology groups, and seventeen individual technologies classified under these groups. The indicators encompass both the extent to which a plant's operations depend on automation, and the amount of investment in automation.

The paper also explores the implications of automation for productivity measurement. A fundamental question is whether more productive plants are also the ones with lower labor share and higher degree of automation. If productivity is Hicks-neutral and the contribution of all inputs is correctly accounted for, automation and the measured productivity residual should be uncorrelated. However, a systematic relationship may be present if automation is correlated with unobserved factors that relevant for output variation. For instance, the use of advanced technology and automation may enhance managerial productivity, inventory management, or coordination on the factory floor – factors that are not captured by standard measures of input usage only. In such cases, a positive relationship between productivity and automation, and a negative one between labor share and automation, may emerge.[7] The empirical results suggest that these relationships indeed hold for the plants in the 1991 SMT.

The findings in this paper contribute to research on the decline in U.S. labor share.[8] One explanation for the decline is the diffusion of labor-saving technologies and automation brought about by the decline in the relative price of capital with respect to labor. This mechanism may be as relevant for manufacturing as it is for the retail, wholesale, and financial sectors, where self-service checkouts, advanced storage systems, automated customer service and other forms of automation have been diffusing.[9] Other explanations include various other factors, such as import intensity and offshoring, the decline in unionization, or labor reallocation.[10] Although the last one of these has received a lot of attention with the rise of productive and large firms ("superstar firms") and the associated increase in industry concentration of employment and sales, the exact mechanisms through which superstar firms emerge, and the role of technology adoption therein, have not been explored in detail. In particular, it is not known to what degree automation and technology use matters for labor share, in addition to the effects of productivity on labor share. This paper provides additional evidence on how productivity and the labor share vary with the intensity of automation across plants.

The analysis in this paper is part of the literature that use the SMT to analyze the connection between technology and plant-level outcomes. Most of the existing work is based on extensive

[7]See Syverson (2011) for a more comprehensive list of factors that, if not properly controlled for, may be systematically related to measured productivity.

[8]For recent examples, see Elsby et al. (2013), Karabarbounis and Newman (2014), Lawrence (2015), Barkai (2016), Autor et al. (2017a,b).

[9]See, for instance, Basker et al. (2017), for an analysis of customer-labor substitution in the context of gasoline stations.

[10]Autor et al. (2013) highlight the role international trade may have on local labor markets. Elsby et al. (2013) argue that the decline of unionization may be considered as factor that depresses wages and reduce employment. Autor et al. (2017a,b) analyze the causes and consequences of labor reallocation.

measures of technology presence.[11] A number of papers look at the relationship between technology presence and plant life-cycle.[12] Others explore the wage premia associated with technology use.[13] The SMT has also been used to study the connection between labor productivity and technology.[14] The analysis differs from previous work in its focus, as the main objective is to estimate plant-level TFP in a way that accounts for the possible connection between input price variation and factor usage and at the same time controls for unobserved productivity differences. For this purpose, intensive measures of investment in automation in the 1991 SMT are more appropriate because they arguably better capture unobserved productivity differences than extensive measures used in earlier studies.[15]

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 documents some stylized facts on the connection between automation and plant characteristics. Section 4 lays out a model of CES production for a manufacturing plant with endogenous technology choice. Based on the model, Section 5 introduces different approaches for estimating TFP. Results and their implications are discussed in Section 6, along with a comparison with standard approaches. This section also explores the connection between the estimated TFP, automation, and production labor share across plants. Section 7 concludes.

2 Data

This section describes the datasets used in the empirical analysis. The main data source on advanced technology and automation is the U.S. Census Bureau's 1991 SMT, part of a collection of surveys on technology use in manufacturing plants conducted in 1988, 1991, and 1993.[16] The

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

[12] Dunne (1994) finds that age and technology use are essentially uncorrelated at the plant level, while Doms et al. (1995) document that capital-intensive plants with advanced technology have higher growth rates and are less likely to fail.

[13] 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. (1997) also examine the connection between wages, skills, and technology using data that connects individual workers to plants. They document that businesses that use a higher number of advanced technologies have more educated workers, employ relatively more managers and pay higher wages. They do not find, however, a significant correlation between skill upgrading and use of advanced technologies at the plant level.

[14] McGuckin et al. (1998) find that establishments that use the most advanced technologies exhibit higher labor productivity than the rest, and that the use of advanced technologies is in general positively related to improved labor productivity performance.

[15] The SMT was conducted for 1988, 1991, and 1993, with extensive measures of technology adoption available in the 1988 and 1993 versions. Some of the plants surveyed in 1988 were dropped and new establishments were added for the 1991 and 1993 SMT. Therefore, the three surveys do not necessarily include the same plants. Despite the differences, some of the previous findings – particularly the relationship between worker wages, labor productivity and technology use – also emerge in the 1991 SMT. This indicates there is some general consistency between the answers in the 1988 and 1993 SMT and the answers to the different questions asked in the 1991 SMT.

[16] During the developmental phase of the survey, the Census Bureau relied on consultations with a broad cross-section of Government, private industry and academic experts. The SMT was partly funded by defense agencies.

survey contains a stratified random sample of about 10,000 observations, representative of nearly 45,000 plants in 1991.

While the SMT pertains to an earlier period, it has several desired features for the type of analysis carried here. First, it contains an exceptionally rich set of measures on the use of automation-related technologies, many of which had already diffused to a large extent even by the time of the survey. In addition, the survey was designed to specifically measure technologies that can substitute for labor, making it ideal for exploring the patterns of capital-labor substitution. It also contains data on a large set of other plant characteristics not available in typical surveys, and can be linked to other Census Bureau surveys to obtain additional plant-level variables. Moreover, the presence of data for the surveyed plants for a long period of time following the survey allows for an analysis of the post-survey evolution of plants with varying degrees of automation.

2.1 Industries

The 1991 SMT has data on 5 major 2-digit SIC manufacturing industries: 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 chosen based on the relatively higher likelihood of reliance on the technologies that are the subject of the survey. They together accounted for about 43% of manufacturing employment around the time of the survey.

The industries in SMT are generally capital intensive, see Table 1. Nevertheless, comparing the production labor share, capital share, capital share-to-labor share ratio, and TFP in SMT industries with the rest of manufacturing industries suggests that they are not special cases in manufacturing. One reason these industries were chosen for the survey may have been the relatively high presence of defense contractors in these industries, which also tend to be more advanced in terms of technology. A number of empirical studies in engineering economics support the view that manufacturing units producing military-use output tend to utilize more advanced technologies.[17] This finding echoes in the SMT: plants that indicate production to military specs have on average higher technology use and investment.[18] Overall, the relatively high prevalence of advanced technologies makes the SMT ideal for exploring the substitution patterns between production labor and capital.

As background information about these industries, Figures 1(a)-1(b) show aggregate labor share measures in the five SMT industries over the period 1958-2007.[19] Both overall and production

[17]See, e.g., Kelley and Watkins (1995,1998,2001).

[18]Question 7 in the 1991 SMT asks plants whether any of the products produced at the plant are manufactured to military specifications.

[19]The industry level data for the SMT industries is obtained from the NBER-CES Manufacturing Industry Database, available at http://www.nber.org/nberces/.

labor share decline in all industries during this period, and the decline dates back at least to the 1970s. Perhaps more surprisingly, capital share also decline in all industries until the mid 1990s, but flatten thereafter – see Figure 1(c).[20] By and large, the trends in labor and capital shares are quite similar across the five industries covered by the SMT over a long horizon. In the year of the survey (1991), the highest production labor share is observed for Fabricated Metal Products (SIC 34) and the lowest for Instruments and Related Products (SIC 38). The highest capital share is in Fabricated Metal Products (SIC 34), and Industrial Machinery and Equipment (SIC 35), and lowest in Transportation Equipment (SIC 37). The ratio of capital share to production labor share, shows slightly different picture, see Figure 2(a). This indicator is the highest in Electronic and Other Electric Equipment (SIC 36), and starts to increase in this industry and in Instruments and Related Products (SIC 38) in the early 1980s, and somewhat later in the remaining industries. Turning to the 5-factor TFP measure, see Figure 2(b), Industrial Machinery and Equipment (SIC 35) and Electronic and Other Electric Equipment (SIC 36) show large increases starting in the early 1990s, whereas other industries experience more modest changes over the entire period.

Overall, these findings suggest that industries in the SMT have largely similar trends in labor and capital shares, but somewhat less so in TFP. Electronic and Other Electric Equipment (SIC 36) stands out as one industry where the trends in capital share, capital share-to-production labor share ratio, and 5-factor TFP are more pronounced post-1990.

2.2 Technologies

The 1991 SMT provides plant-level intensive measures of technology adoption, use, and investment for four broad technology types, which include 17 individual technologies, listed in Table 2. Some of the technologies (e.g. Robots, Automated Storage and Retrieval Systems, Automated Guided Vehicle Systems, and Automated Sensor Based Inspection/Testing Equipment) are directly aimed at automating tasks performed by labor, whereas others (e.g. Computer Aided Design/Engineering, Computer Aided Manufacturing, Local Area Networks) can either facilitate or support automation of tasks. The analysis treats the technologies as parts of automation in a plant. All technologies have the potential to replace workers engaged in production. [21]

The same technologies are the subject of the survey questions in all three waves of the SMT – 1988, 1991, 1993. The 1991 SMT is the key input for this paper because of its specific questions on intensive measures, such as the amount of past and planned future investment in advanced technologies. However, the 1991 SMT does not provide information on which of the specific 17 technologies were present at the time of the survey. This information is instead available for the

[20] Note that the capital stock measure is not quality-adjusted.

[21] Identifying which technologies matter most for automating tasks and replacing labor is of importance – a challenge left for future work.

plants surveyed in the 1988 and 1993 versions of the SMT. Table 2 shows the rate of diffusion across plants based on these two surveys. While robots are relatively less common in U.S. manufacturing during these survey years, many other technologies, such as numerically controlled/computer-numerically controlled (NC/CNC) machines, computer-aided design, engineering and manufacturing, programmable controllers, computer networks, sensor-based inspection, and flexible manufacturing cells/systems, have relatively high diffusion rates. Given that the relative diffusion rates of the technologies are highly similar in the 1988 and 1993 SMT, the variation in diffusion rates across the technologies is likely similar for the 1991 SMT.

2.3 Measures of Technology Use and Investment

The specific intensive measures of technology used in this paper are based on four main check-box-type questions, described in Table 3. The questions ask about current and future dependence of operations on technology, as well as about past and future investment in technology. The questions were asked for each of the four broad technology types, providing a rich characterization of adoption and use of various technologies by the plant. For the purposes of this study, each response was recoded into a numerical category, see Table 3 for more details.[22]

An important advantage of the 1991 SMT with respect to the 1988 and 1993 SMT is the more accurate measurement of the contribution by advanced technology. The dependence of operations on technology and the dollar-value of investment in technologies arguably better reflect technology-dependence than an indicator of whether the plant has any specific technology, or how many of the technologies it uses – measures available in the 1988 and 1993 SMT. For example, while two plants may both have robots, a larger dollar value of investment in robots in the first plant compared to the second better captures the fact that the first one relies more heavily on robots and therefore automation may have a larger effect on the plant's operations and workforce. These considerations are important for the purposes of this paper because identification is based on cross section variation in these indicators.

A caveat on measurement is that the responses are recorded as ordinal values. While the ordinal scale may introduce noise, it allows an ordering of plants' technology usage and investment intensities across different technology types, implying that the cross section variation in these measures can be used for identification.

[22]It is important to emphasize that although higher categories indicate a higher use or investment level, higher categories do not correspond to a linear increase in responses.

2.4 Other Plant Characteristics

In addition to technology indicators, the 1991 wave of the SMT contains a variety of measures on plant characteristics, including employment, value of shipments, age, export intensity, the presence of a union contract for production workers, the average price of plant's products, production for military purposes, and government contracting/subcontracting.[23] As in the case of the questions related to technology, these measures are only available in categorical or ordinal form.

Continuous measures of output and inputs were obtained from the 1991 Annual Survey of Manufactures (ASM) for most of the plants surveyed in the 1991 SMT. Since the SMT was conducted separately from the 1991 ASM, its sampling frame is different from that of ASM and some SMT plants are not in the 1991 ASM. The 1990 ASM and 1992 Census of Manufacturers (CMF) were used to supplement some of the continuous variables.[24]

For productivity measurement, data on input usage and prices are required. However, the unavailability of such data for many plants leaves a smaller sample for productivity analysis.[25] Data from ASM and CM is used to construct plant-level inputs and output. The measure of output is the deflated value of total value of shipments. Production labor input is measured by production worker hours and the ratio of the total wage bill to production worker wages. 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. The intermediate input variable is obtained as the sum of cost of parts, contracted work and goods resold. The energy input is composed of costs of electricity. 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. Plants' initial capital stock to a deflated book value taken from the ASM and CM.

A separate unbalanced panel of ASM plants is also utilized in the estimation of the elasticity of substitution between production labor and capital. This dataset uses plants in the ASM for the period 1987-1996 for industries that are covered by SMT. 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 plants surviving till 1997 and 2002 are used to study the evolution of labor share within the next 5 to 10 years as a function of technological sophistication as of 1991.[26]

[23]Some of these measures (e.g. unionization, export intensity and production for military) provide rare opportunities to explore relatively less known properties of manufacturing plants. For instance, the responses about the presence of a union contract can be used to assess the relationship between unionization and other plant characteristics. For a use of the survey for this purpose, see Dinlersoz et al. (2017).

[24]If a plant was not found in 1991 ASM, 1992 CMF was searched for this plant. If found, the values of the continuous variables reported in the 1992 CMF were used. If a plant was neither in 1991 ASM and 1992 CMF, 1990 ASM was used to attach values to the continuous variables for the plants that appeared in 1990 ASM.

[25]Any administrative records cases were also dropped.

[26]The unavailability of the type of data collected in 1991 SMT for other years prevents a full dynamic analysis

3 Stylized Facts on Automation and Labor Share

In this section, the basic facts about the relationship between technology adoption and plant characteristics are laid out, focusing on capital and labor usage. The degree of automation is measured using a technology index based on information about four technology types, see Table 2.[27] The index averages re-coded plant-level responses to four questions – listed in Table 3 – about technology dependence and investment across these technologies. The index spans continuous values between 0 and 5, each value representing an average of past, current, and planned future technology investment and use intensity.[28] An index value of zero indicates that the plant has virtually no reliance on automation. Higher values indicate greater use of, and investment in, automation.

3.1 Labor Share and Automation across Plants

The subjective assessments by plants in the 1991 SMT indicate that labor cost reduction is deemed an important benefit from the use of advanced technologies related to automation, as shown in Figure 3.[29] Labor cost reduction comes as the second most common benefit cited by plants next to quality improvement, followed by increase in flexibility of plant's production. While the distribution of responses suggests that a major motivation for using automation is reduction of labor costs, a quantitative assessment of labor cost relative to the plant's revenues is not possible based on the data collected in the survey alone. To that end, measures of labor share obtained from ASM and CM are related to measures of the degree of automation.

Figure 4 plots non-parametric local polynomial smoothing estimates of labor's share in a plant's total value of shipments, as a function of the technology index. Pointwise 95% confidence intervals are shown as dotted lines in the figure. Three key observations can be made from Figure 4.

First, labor share is lower for more technologically advanced plants: it drops from 29% of revenues to 24% (a decline of 17%) as the technology index increases. The decline is statistically significant for much of the index range.[30] Second, the decline in labor share is driven by the decline in production labor share rather than non-production labor share. The former drops nearly by half, from 17% to 9% (see Figure 4(b)) and this decline is statistically significant, whereas the

of the evolution of automation intensity at the plant level.

[27] There is also an additional question, not listed in Table 3, about the expected investment associated with future plans on technology adoption/upgrade, i.e. expected cost of future acquisitions (Question 13 in the 1991 SMT Report Form). Incorporating this question to the technology index makes little difference in our results and conlcusions.

[28] Alternative measures are also considered, as discussed below.

[29] The responses in Figure 3 are to Question 10 in the survey: "What benefits have you derived from the use of technologically advanced equipment or software in this plant?". The response category "Not Applicable" is omitted in the figure.

[30] The confidence intervals get larger towards the high end of the technology spectrum owing to the relatively small sample of plants in that region and the one-sided nature of the kernel smoothing near the end of the sample.

later actually increases slightly from 12% to 14% (see Figure 4(c)), though the increase is not statistically significant. These two observations suggest that technologically advanced plants tend to have much lower fraction of their revenues dedicated to compensating production labor, but a slightly higher fraction to non-production labor.

Third, plants with higher levels of automation also tend to have a lower fraction of their workforce engaged in production (Figure 4(d)). At the lowest levels of the technology index, the fraction of production workers in plant employment is about 70%. This fraction drops sharply to nearly 50% at the highest levels of the index – a decline of almost 30%. Figure 5 provides additional evidence on the relative input usage. The capital share in the value of shipments increases with the technology index (Figure 5(a)). As a result, the ratio of capital share to labor share is also an increasing function of the degree of automation in a plant as measured by the technology index (Figure 5(b)). The same statement holds for the capital-labor ratio, plotted in Figure 5(c).[31]

Technology is also closely related to other measures of plant performance (Figures 6-8). Namely, labor productivity increases with technology index, especially in the case of production labor (Figure 6), a finding robust to alternative ways to measure labor productivity (Figure 7).[32] In addition, the average wage bill per worker increases for both types of labor as the technology index increases (Figures 8).

The relationships between the technology index and plant-level outcomes are robust to other controls such as plant size and age, unionization, or whether or not a plant exports. Tables 4-6 show the estimated coefficients of the technology index conditional on these controls.[33] Tables 4-6 also feature, for robustness, an alternative technology index that only includes the average investment indicator across the four technology groups based on survey question 2 in Table 3. The results indicate that plants that rely more on, or invest more in, technology, tend to have lower production labor share and exhibit higher production labor productivity and average wage. A 1% increase in technology index is associated with a 0.04-0.08% decline in production labor

[31]The patterns in Figures 4 and 5 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.

[32]Note that labor's revenue share, LS, can be written as $LS = \frac{wl}{r} = w\left(\frac{r}{l}\right)^{-1} = w \times (LP)^{-1}$, where w is average wage, l is employment, r is revenue, and LP is revenue productivity of labor. Hence, labor share is inversely related to labor productivity, and positively associated with average wage.

[33]These characteristics include five plant size (employment) categories (1-20 emp, 20-99 emp, 100-499 emp, 500-999 emp, 1000+ emp), four 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 righrs 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. All dependent variables are expressed in logarithms, and an inverse hyperbolic sine transformation is used for the technology index. The transformation allows observations with technology index value of zero to be kept in the analysis. Hence, the estimated coefficients can be interpreted approximately as elasticities.

share, 0.12-0.14% increase in production labor productivity, and 0.08-0.09% increase in average production worker wage. 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.[34] These results confirm the bivariate relationships discussed earlier.

While not included in Tables 4-6, some of the plant-level controls are also significantly related to production labor share.[35] For example, older, foreign-owned, and unionized plants have higher production labor share, while larger plants and plants that export more than 50% of their products have lower production labor share. The patterns in Table 4-6 are also robust, and even more pronounced in many cases, when value added is used as an alternative to total value of shipments in calculating labor share and labor productivity measures.

3.2 Change in Labor Share and Automation

The measures of automation are available for 1991 only, so a complete panel analysis that considers changes in the degree of automation is not possible. Instead, the approach is to analyze post-1991 evolution of plants that likely depend on the degree of automation, and explore the 5- and 10-year changes in key outcomes as a function of the technological sophistication of the plant as of 1991. This approach may be informative about the dynamic relationship between automation and factor usage because several automation-related technologies, such as computer aided manufacturing and local area networks, are likely to remain in place over time.[36]

On average, the data indicates that production labor share declines in surviving plants over time. The change in production labor share over time, however, is not uniform across plants. While many plants experience negative growth rates in production labor share, some experience a positive one. To explore the connection between automation and change in labor share, the following specification is estimated

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

[34] Non-production worker category includes labor with various education and skill levels, and this composition effect may be hiding the potentially divergent patterns for different worker types classified in the group. Census Bureau defines a non-production worker as a worker engaged in the following activities: factory supervision above the working foreman level, sales, sales delivery, advertising, credit collection, installation and servicing of own products, clerical and routine office functions, executive, purchasing, financing, legal, professional, and technical.

[35] The estimated coefficients of these variables are not released to reduce the amount of information disclosed about the sample of plants studied.

[36] It would be possible to use the 1988 and 1993 surveys to analyze the effect of a change in the degree of automation on the change in plant outcomes. However, such an approach has several drawbacks. First, there is significant attrition between the 1988 and 1993 waves of the surveys. Second, the technology indices in these surveys are extensive measures. Third, prior research with these two surveys indicate some recall bias.

where ΔY_i is the log difference in the labor share between 1991 and 1997, or between 1991 and 2002, ΔE_i is the log difference in total plant employment over the same horizon, and I_i is the technology index as of 1991. X_i includes other plant-level controls and industry effects as in Tables 4-6. ΔE_i controls growth-related heterogeneity.[37] 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. The specification in (1) is also implemented using the log difference in production labor productivity as the dependent variable.

The results in Tables 7 and 8 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. Specifically, 1% higher technology usage or investment in 1991 is associated with 0.07-0.08 percentage point lower labor share growth.[38] The effect of higher automation in 1991 on labor productivity growth is the opposite. A 1% higher technology index in 1991 is associated with 0.07-0.11 percentage point higher growth in production labor productivity. In addition, employment growth has a negative association with labor productivity.[39] Controlling for survival bias using a Heckman correction confirms these conclusions: Tables A2 and A3 in Appendix A.2 show qualitatively similar results. Conclusions are also stronger when value added is used to measure production labor share and productivity.

Overall, the stylized facts indicate that automation is tied to labor usage and labor productivity in a statistically and economically significant way, both across plants and over time. Models of production that yield constant labor share across plants or over time, such as Cobb-Douglas technology, cannot appropriately account for the facts documented. The systematic differences across plants in factor usage and technology investment can be better captured by models where plants choose and alter the degree of automation.

4 The Model

This section offers a model of plant-level production that can account for the stylized facts on capital-labor substitution presented above. A key feature of the model is that a plant adjusts its capital-labor ratio in response to changes in the relative price of these inputs. The other important feature is that the nature of the relationship between the capital-labor ratio and relative price is fully determined by the degree of substitutability between these inputs.

[37] Growing plants that hire more employees are expected to experience a rise in labor share.

[38] These results are conditional on overall employment growth: the estimates show that a 1 percentage point increase in employment growth is associated with a 0.11-0.17 percentage point rise in the growth rate of production labor share.

[39] A 1 percentage point increase in employment growth is associated with a 0.10-0.17 percentage points decline in production labor productivity growth.

4.1 Technology

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, L_n is non-production labor, M and E are materials and energy, K denotes capital, and L_p is production labor. Freely variable inputs L_n, M, and E are combined using a Cobb-Douglas aggregator with parameters $0 < \beta_j$, $\forall j$.[40] Quasi-fixed inputs K and L_p, are aggregated using a CES form into a composite input, $T_i = [\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}]^{1/\rho}$. The assumption of quasi fixity of these two inputs is justified if K and L_p are subject to non-linearities (see Caballero et al. (1997)) or non-convex adjustment costs (see Cooper and Haltiwanger (2006) and Bloom (2009)). While quasi-fixity of capital is easier to motivate, the case of production labor might be less so. However, the presence of unions for production workers (especially in the data studied here), and general labor adjustment costs associated with hiring and firing, implies that there are likely substantial non-convex adjustment costs for this type of labor.[41]

The parameter $\rho \in \mathbb{R}$ determines the elasticity of substitution, σ, between production labor and capital – ρ and σ are related as $\sigma = 1/(1-\rho)$. The variable $\alpha_i \in (0,1)$ in T_i is referred to as the technology of the plant. It is a decision variable, and the plant sets T_i by adjusting α_i in response to changes in the relative price of K and L_p. Allowing for α_i to be endogenous is a deviation from most of the earlier work because α_i is generally assumed to be an exogenously given constant within an industry, see, among many others, Lawrence (2015) and Raval (2017). In a fully specified model, Hicks-neutrality implies that α_i determines output only through its effect on the plant's composite input. In other words, the production function in (2) does not impose any restriction on the relationship between productivity and other plant characteristics.[42]

Note also that α_i is related to, but different from, factor-augmenting shifters used in some CES representations. While it is possible to rewrite the CES component of (2) to fit standard forms

[40] The Cobb-Douglas assumption for this part of the production function is mainly a simplification, since the main focus of this paper is on understanding the connection specifically between production labor and capital – in particular, capital in the form of advanced technology. It is possible to extend the analysis by using nested CES specifications that can allow for varying degrees of substitutability between both labor inputs and capital, as well as energy and materials.

[41] Non-production labor may be subject to similar considerations. For instance, the ranks of R&D workers and managers embed valuable organization capital and may be slow to adjust. Nevertheless, the rest of the non-production workers, including office personnel, administration, and plant maintanence and cleaning staff, may be less so. Overall, the assumption is that non-production labor is, on average, less subject to adjustment frictions.

[42] If model (2) is not fully specified, a correlation between productivity and other plant characteristics may emerge. For instance, if $\theta_i \equiv \theta(\alpha_i)$, the choice of α_i directly affects productivity in addition to its effect through the composite input. Such an assumption is appropriate if the adoption of labor-saving technologies results in more flexibile production, improves coordination of production processes, or allows management to be more effective in monitoring production. All these mechanisms would yield positive correlation between α_i and θ_i.

with factor-augmenting technical change, the key difference relative to those approaches is that α_i is chosen by the plant in response to movements in relative input prices over time or across plants, instead of being taken as an exogenous technology parameter as in the case of factor-augmenting shifters.[43]

Standard functional forms are limiting cases of equation (2). For example, Cobb-Douglas technology is obtained as $\lim_{\sigma\to 1} T_i$. Leontief and linear technologies are given as $\lim_{\sigma\to 0} T_i$ and $\lim_{\sigma\to\pm\infty} T_i$, respectively.[44] The specification in equation (2) is different from standard models of capital embodied technical change. While a higher level of α_i embodies more capital in the composite input, this is the result of plants' endogenous technology choice in response to price changes, not of exogenous productivity shocks, as would be the case in a standard model of capital or labor embodied change – see also Acemoglu and Restrepo (2018a,b) for an assessment of modeling automation as exogenous capital or labor augmenting technological change, which have implications on equilibrium labor share and wages that do not necessarily line up with the accumulated evidence.

4.2 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.

4.2.1 Exogenous Output Prices

Plants produce a homogenous good with its price fixed and normalized to one. All factor prices are allowed to vary across plants, as opposed to the typical assumption that they are constant. The assumption of heterogenous input prices is justified if there are differences across plants in terms of the quality of their inputs. One example would be the case in point, i.e. where plant-level capital stocks differ in the extent to which they contain automation-related technologies.[45] The first-order conditions imply that the capital-to-production labor ratio, and the relative weight on

[43] With factor-augmenting shifters, the CES production function can be written as $[(AK_i)^\rho + (BL_{pi})^\rho]^{\frac{\gamma}{\rho}}$. Setting $A = \alpha^{2(1-\rho)/\rho}$ and $B = (1-\alpha)^{2(1-\rho)/\rho}$ gives the CES component of (2).

[44] It is straightforward to generalize T_i. For example, $T_i = [\alpha_i^{\zeta/\sigma} K_i^\rho + (1-\alpha_i)^{\zeta/\sigma} L_{pi}^\rho]^{1/\rho}$, where $0 < \zeta$. Equation (2) then corresponds to the analytically convenient case $\zeta = 2$. This normalization can be motivated by noting that the actual value of ζ has implications for σ. In the model, more technology means arguably more complementarities between capital and labor, which in turn would be encapsulated with lower σ. Given that CD technology is a special case of CES with σ=1, these considerations imply $\sigma < 1$ in the SMT if the true data generating process is CES. Choosing $\zeta = 2$ restricts the parameter space but does not necessarily restrict the estimate of σ to be less than 1, and therefore seems general enough.

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

capital and production labor, can be written as[46]

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

$$\frac{\alpha_i}{1-\alpha_i} = \left(\frac{w_{pi}}{w_{ki}}\right)^{1-\sigma} . \tag{4}$$

These expressions highlight the key data generating mechanism of the model: both α_i and the capital-labor ratio are tied to relative input price variation and the nature of these relationships is fully determined by σ. Equations (3) and (4) together imply that $\frac{K}{L_p}$ and $\frac{\alpha_i}{1-\alpha_i}$ are increasing in the relative price of production labor, as long as $\sigma \in (0,1)$. An increase in the relative price of production labor induces the plant to substitute away from L_p by increasing K. Solving equation (4) for the weight of production labor in T_i yields $1-\alpha_i = \left(1 + (w_{pi}/w_{ki})^{1-\sigma}\right)^{-1}$, implying that $1-\alpha_i$ is decreasing in the relative price of production labor when $\sigma \in (0,1)$. An implication is that, if the true data generating process lines up with a CES specification that implies capital-production labor substitution, the estimates of σ should be less than one.

It is important to note that (3) indicates a convex relationship between capital-production labor ratio and relative price when $\sigma \in (0,1)$. There is also a convex relationship between capital-production labor ratio and technology as measured by $\frac{\alpha_i}{1-\alpha_i}$

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

since $\frac{2-\sigma}{1-\sigma} > 1$ under $\sigma \in (0,1)$. This is in contrast to the case of standard Cobb-Douglas specification, where both of these relationships are linear.[47] The relationships are instead concave in the case of a standard CES formulation without endogenous technology, when $\sigma \in (0,1)$.[48] The empirical analysis suggests a convex relationship between technology index and capital-production labor ratio (see Figure 5(c)). The empirical patterns hence line up better with a convex relationship in (5).

Since shares of input expenditures are of primary interest, it is useful to describe their properties using the first-order conditions. Combining equations (3)-(4) yields an expression for the share of

[46]Cost minimization implies 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 $TC_i = \sum_j w_{ji}X_{ji} = \lambda^* Q_i \left(\sum_j \beta_j + \gamma\right)$.

[47]Under Cobb-Douglas, $\frac{K_i}{L_{pi}} = \frac{\alpha_K}{\alpha_L}\frac{w_{pi}}{w_{ki}}$. Hence, the capital-labor ratio is linear in relative price, given technology. It is also linear in technology ($\frac{\alpha_K}{\alpha_L}$), given relative price.

[48]Under CES given by $[\alpha_i K_i^{\rho} + (1-\alpha_i) L_{pi}^{\rho}]^{\gamma/\rho}$, one obtains $\frac{K_i}{L_{pi}} = \left(\frac{\alpha_i}{1-\alpha_i}\right)^{\sigma} \left(\frac{w_{pi}}{w_{ki}}\right)^{\sigma}$. Hence, when $\sigma \in (0,1)$, the capital-labor ratio is concave in relative price given technology, and concave in technology $\frac{\alpha_i}{1-\alpha_i}$, given relative price.

production labor in the cost of T_i

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

That is, the optimal weight of L_{pi} in T_i is also its share in the cost of T_i. Since the revenue share of T_i equals γ, the revenue share of production and total labor can be written, respectively, as

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

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

Equations (6), (7) and (8) indicate that labor share measures are decreasing in α_i. The rate at which they decrease is captured by their sensitivity to α_i.[49]

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 $\sigma \in (0, 1)$. One difference relative to the results for Cobb-Douglas technology is that imposing constant returns to scale (CRS) is not sufficient in order to identify factor elasticities. Although variable input elasticities are identified by cost shares under CRS, the share of the composite input, T_i, in total costs underestimates the contribution of γ to returns-to-scale, irrespective of the value of returns-to-scale.[50]

When returns-to-scale is a free parameter, 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 $\beta_j = \frac{w_{ji}X_{ji}}{Q_i}$, and $\gamma = \frac{w_{ki}K_i}{Q_i} + \frac{w_{pi}L_{pi}}{Q_i}$, 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.[51]

[49] If $\gamma < 1$, the rate of decline in (7)-(8) as α_i increases is smaller in absolute value than the rate of decline in (6). When $\gamma > 1$ the relationship is reversed. When $\gamma = 1$, all three shares decline at the same rate.

[50] The results under Cobb-Douglas carry over to variable input elasticities: under increasing (decreasing) returns to scale, cost shares of variable input expenditures underestimate (overestimate) the factor elasticities.

[51] The corresponding condition for variable input X_j can be written as $w_{ji} = \beta_j \frac{Q_i}{X_{ji}}$. For K_i and L_{pi} these read $\frac{Q_i}{K_i}\gamma\alpha_i^{\frac{2}{\sigma}} K_i^{\rho} T_i^{\gamma-1} = w_{ki}$, and $\frac{Q_i}{L_{pi}}\gamma(1-\alpha_i)^{\frac{2}{\sigma}} L_{pi}^{\rho} T_i^{\gamma-1} = w_{pi}$. These first order conditions reveal that the model is consistent with the following timing. Quasi-fixity implies that K_i and L_{pi} are pre-determined at the beginning of period t. Since α_i and quasi-fixed inputs are jointly determined, as the previous conditions show, α_i is also pre-determined. Next, θ_i is realized and variable inputs are chosen.

4.2.2 Isoelastic Residual Demand

The previous section imposed price taking behavior in output market. An alternative to fixed output prices is to postulate that the plant's residual demand is isoelastic.[52] 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 isoelastic demand marginal revenue products are smaller than marginal products. To see this formally, let R_i denote plant-level revenues P_iQ_i, and write the first order conditions for the jth variable input and the quasi-fixed inputs as

$$\begin{aligned} \frac{w_{ji}X_{ji}}{R_i} &= \kappa\beta_j \\ \frac{w_{ki}K_i}{R_i} + \frac{w_{pi}L_{pi}}{R_i} &= \kappa\gamma, \end{aligned} \tag{9}$$

where the second line combines the conditions for K and L_p. The implications of these conditions for the relationship between w_{pi}/w_{ki}, K_i/L_{pi} and $\alpha_i/(1-\alpha_i)$ are the same as in equations (3)-(4). Intuitively, since demand affects all inputs, their relative allocations do not change in the wake of a demand shock. An important difference relative to Section 4.2.1 is that factor elasticities depend on both the revenue share of input expenditures and the inverse of the demand parameter κ. Therefore, under $\kappa \in (0,1)$ revenue shares underestimate variable factor elasticities and γ.[53] 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.[54] Preliminary analysis indicates that this price information has no additional explanatory power conditional on continuous variables such as capital and labor.

5 Semi-parametric Estimation

The estimation strategy follows a structural approach. First, σ is estimated by transforming (3) into an estimable equation where plant-level variation in production labor is projected onto cross-sectional differences in plant-level production wages and capital. Under the assumptions of the model, this projection is informative about the substitution patterns between K and L and

[52]This approach is commonly used in the literature. Recent examples include De Loecker (2011), Bartelsman et al. (2013), Foster et al. (2016, 2017), and Haltiwanger and Wolf (2018).

[53]Solving the first order conditions for the elasticities yields $\beta_j = \kappa^{-1}\frac{w_{ji}X_{ji}}{R_i}$, and $\gamma = \kappa^{-1}\left(\frac{w_{ji}X_{ji}}{R_i} + \frac{w_{pi}L_{pi}}{R_i}\right)$.

[54]The categorical price variable measures average price for the products of a plant and is available in the 1988 and 1992 SMT only. Therefore this information is not available for all plants in the 1991 SMT.

therefore can be used to identify σ. The other parameters of production function (2) are estimated conditional on the estimate $\widehat{\sigma}$, using a modified version of the approach described in Haltiwanger and Wolf (2018). The remaining coefficients are determined conditional on these parameters.

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 *i.i.d.* error term. Given data on l_{pi}, k_i, and their prices, this equation can be estimated by running the regression

$$l_{pi} = \delta_1 \ln w_{pi} + \delta_2 \ln w_{ki} + \delta_3 k_i + u_i. \tag{10}$$

The wage rate for production labor, w_{pi}, is obtained by dividing production labor costs by production worker hours. This approach implies that OLS estimates of δ_1 are affected by division-bias, which is addressed using geographic variation in wages, where w_{pi} is instrumented using a state- and 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 of capital w_{ki} is not observed in the data. Capital costs are calculated by combining industry-specific rental prices and plant-level capital measures.[55] This approach results in measures of $w_{ki}K_i$ and K_i, but not w_{ki}, implying that δ_2 is not identified. Under the assumption that w_{ki} is plant-specific, its effect is accounted for by a plant-level fixed effect, in which case δ_1 and δ_3 are identified in the first-differenced version of (10). This approach is justified when plant-level capital prices are persistent, for instance when they follow a random walk.

Given an estimate of δ_1, $\widehat{\sigma}$ can be obtained using $\widehat{\sigma} = \widehat{\delta}_1 + 2$. Note that this estimate is influenced by the choice of $\zeta = 2$ in (2). This choice is driven by analytical convenience and its ability to generate plausible values for σ. In the sample of plants studies here the use of automation and advanced technology suggests that capital and production labor are likely less substitutable than what a Cobb-Douglas specification would imply. In other words, one expects plausible estimates of σ to fall in the range $(0, 1)$. Choosing $\zeta = 2$ does not preclude this possibility and also allows for a simpler analytical framework.

[55] See Foster et al. (2016) about the properties of this data.

5.2 Factor Elasticities

The estimation strategy for factor elasticities builds 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 provides variables that record categorical responses on how much the plant invested in four technology types in the previous three years – see question 2 in Table 3. Although the variables are categorical, they provide direct information on cross-plant differences in technology investment. The responses are combined 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 the majority of earlier studies that mostly rely on general investment to control for unobserved productivity differences during estimation.[56]

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 estimate the elasticities of freely variable inputs. Given 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 identify factor elasticities without projecting revenue variation on proxies, state variables, or variable inputs.[57] This feature 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 to 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. Below is an outline of the estimation approach:

1. Obtain IV estimates of σ based on (10): $\widehat{\sigma} = \widehat{\delta}_1 + 2$.
2. Compute $w_{ji}X_{ji}/R_i$, and estimate input elasticities using average revenue share of the input, $\widehat{\beta}_j = 1/N \sum_i w_{ji}X_{ji}/R_i$. Averaging mitigates the effects of measurement error, and is often used in empirical productivity literature.
3. Net out the contribution of variable input costs from revenue to obtain $\widehat{\mathcal{B}}_i = r_i - \sum_j \widehat{\beta}_{j_i} w_{ji} X_{ji}$.

[56] 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).

[57] See Section 4.2.2 for more details.

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 contribution of the composite input
$$\widehat{T}_i \equiv \frac{1}{\widehat{\rho}} \ln \left[\alpha_i^{2/\widehat{\sigma}} K_i^{\widehat{\rho}} + (1-\alpha_i)^{2/\widehat{\sigma}} L_{pi}^{\widehat{\rho}} \right]. \tag{11}$$

5. Following Haltiwanger and Wolf (2018), determine the joint contribution of state variables and the proxy by estimating
$$\widehat{\mathcal{B}}_i = \phi(Z_i, p) + v_i, \tag{12}$$
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. State variables include, but are not limited to, $\widehat{T}_i$ and other plant characteristics, such as plant age. If the only state variable is $\widehat{T}_i$ and if technology investment can be subsumed into a single indicator $\bar{t}_i$ then $Z_i = \left(\widehat{T}_i, \bar{t}_i\right)'$.[58]

6. Given fitted values $\widehat{\phi}_{it}$ from equation (12), use nonlinear least squares to estimate
$$\widehat{\mathcal{B}}_{it} = \delta_T \widehat{T}_{it} + h\left(\widehat{\phi}_{it-1} - \delta_T \widehat{T}_{it-1}\right) + \nu_{it}. \tag{13}$$
where h is a second-order polynomial in its argument. Under the assumptions underlying equations (2) and (11), δ_T, the coefficient of $\widehat{T}_i$, in regression (13) identifies γ.

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 – see Table 3. The responses of each plant in 1991 are averaged over the four technology groups ($\bar{t}_i$), to determine $\widehat{\phi}_i$ in (12), which is the joint contribution by $\widehat{T}_i$, $\bar{t}_i$ and plant age. Under the assumptions of the model, this value can be used to control for unobserved productivity differences across plants when estimating δ_T using data from 1992 in (13). If the plant-level productivity process is Markovian –a standard assumption in the empirical productivity literature– then δ_T is consistently estimated in regression (13). The standard error of $\widehat{\delta}_T$ is estimated using a bootstrap approach, because $\widehat{\delta}_T$'s distribution is non-standard.

6 Results

6.1 Elasticity of Substitution

The estimates of σ vary between 0.38 and 0.71 depending on the methods used, see Table 9. These estimates are less than one and fall to a range consistent with what recent work found using

[58] 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 T_i if the only unobservable is productivity and if investment in technology is an increasing function of productivity.

similar Census data.[59] 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.

An estimated value of σ that is less than one implies that capital and production labor is less substitutable than what a Cobb-Douglas (CD) specification would imply. In other words, the isoquants of the CES production function has more curvature than that of CD. This lower substitutability is plausible, considering the fact that many plants in the sample rely on automation, and hence their capital stock is more likely to represent advanced technology that may not be easily exchanged with labor. For example, removing a robot from the production process may require increasing production labor a lot more to maintain the same level of production.

Given that the estimates of σ are less than one, the relative weight of capital, $\frac{\alpha_i}{1-\alpha_i}$, is increasing in relative price of production labor under the assumptions of the model. The baseline $\widehat{\sigma}$, shown in column 1, is determined by a cross-sectional IV. Column 2 contains the results of estimating (10) without k_i as an explanatory variable. This approach may be justified if wage and labor are better measured than capital, arguably the case for ASM and CM. These two surveys collect 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. If capital is measured with error then it is a priori unclear whether including k_i in (10) is useful. The similarity of the σ estimates suggests that production labor data alone is informative for substitution patterns. Columns 3 and 4 show additional robustness checks, where (10) is estimated using ASM data on all plants in SMT industries between 1987 and 1996. Instead of using geographic wage variation as an instrument, these calculations are based on lagged differences of plant-level wages as instruments in a GMM framework, see Arellano and Bond (1991). The GMM estimator yields similar $\widehat{\sigma}$s.

6.2 Factor Elasticities

Table 10 shows estimated factor elasticities conditional on the baseline $\widehat{\sigma}$. As the variation in column $\widehat{\gamma}$ indicates, all reviewed methods yield comparable $\widehat{\gamma}$s, suggesting that the contribution of the composite input to returns-to-scale is between 0.17 and 0.25, 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 significantly less than one, which may be surprising at first sight. However, under isoelastic demand these point estimates are revenue elasticities implying that they can be considered as lower bounds for factor elasticities

[59] It would be misleading to use standard Wald- or χ^2-type distributions to test $\mathrm{H}_0 : \sigma = 1$ because the Cobb-Douglas structure is a limiting case in equation (2). This means that a standard test would be unlikely to able to differentiate between H_0 and a case that is arbitrarily close to the limit, which is precisely what would be required in the present context. The deviation from H_0 is a necessary modeling decision if one wants to account for certain properties of the data. Statistical tests in the present context were proposed by Kmenta (1967) and Brown (1970).

– see Section 4.2.2 for more details.

6.3 Properties of the TFP Estimates

This section investigates the empirical implications of the modeling assumptions discussed in section 4. For the sake of robustness, two commonly used Cobb-Douglas productivity measures are evaluated against three productivity measures that are based on the CES specification. The first standard measure, denoted by $\mathrm{CD}_{\mathrm{CRS}}$, is derived under the assumption of constant returns-to-scale and Cobb-Douglas technology. The second productivity indicator, labeled as $\mathrm{CD}_{\mathrm{NCRS}}$, is calculated under the assumption of non-constant returns-to-scale, homogenous products, and price taking behavior.[60] The CES productivity indices correspond to the three specifications shown in Table 10. The first of these, denoted by $\mathrm{CES}_{\mathrm{FOC}}$, is based on $\widehat{\gamma}$ obtained as a plant-level average of the first-order condition (9) under the assumption of exogenous and homogenous output prices and endogenous technology choice. The second CES measure is simply a variant of the first one, in the sense that it is calculated under the same assumptions, but γ is estimated using nonlinear least squares. This specification is labeled as $\mathrm{CES}_{\mathrm{EN}}$. The third one is similar to the second one except that $\alpha_i = 1/2$ is imposed. This exogenous technology specification is denoted by $\mathrm{CES}_{\mathrm{EX}}$.

The descriptive statistics, shown in Table 11, 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. The CD approach yields more observations around the mode and in the tails, indicated by higher kurtosis and lower dispersion. Bivariate correlations echo these differences, as shown in Table 12. 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 productivity distributions obtained under alternative technologies are systematically different. Given that the SMT collected data from industries where automation is likely to substitute for labor, the distinction between CD and CES technologies is expected to be relevant. The results of Kolmogorov-Smirnov tests in Table 13 confirm this conjecture: p-values indicate that CD- and CES-based 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 technology 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 Kolmogorov-Smirnov test is useful to assess whether two distributions can be considered different in the statistical sense. However, it is not informative about possible sources of the differ-

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

ence. In order to shed some light on the nature of the differences discussed above, Appendix A.1 provides a detailed decomposition in which the difference between CD_{CRS} and CES_{EN} productivity residuals 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}. It is a useful metric because it helps understand why the KS test rejects the null of equivalence. The difference can be written as

$$\Delta_i = \frac{\widehat{\gamma}}{\widehat{\rho}} \ln \left[\alpha_i^{2/\widehat{\sigma}} K_i^{\widehat{\rho}} + (1-\alpha_i)^{2/\widehat{\sigma}} L_i^{\widehat{\rho}} \right] - \left(\widehat{\beta}_k \ln K_i + \widehat{\beta}_l \ln L_{pi} \right). \tag{14}$$

Interpreted as a sample statistic, Δ_i accounts for all the specification error if the estimation error is the same under the two specifications, because in this case it is the only component that contributes to the difference between CD_{CRS} and CES_{EN} productivity. Appendix A.1 explores the properties of Δ_i in more detail. The results of evaluating (14) in the sample of plants used for productivity estimation suggest that Δ_i is negative for the majority of plants. This means the CD_{CRS} input index (the second term in 14) is systematically higher than the CES_{EN} input index (the first term in 14). In other words, CD_{CRS} tends to underestimate productivity if the true underlying productivity is CES_{EN}. In addition, 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 The Nexus of Automation, Productivity, and Labor Share

In order to shed light on the relationship between productivity and automation, Table 14 reports the results from regressing the share of production labor costs in revenues on the estimated TFP (CES_{EN}) and technology use and investment, controlling for other plant characteristics. The main message of Table 14 is that the revenue share of production labor costs is lower in more productive and automated plants. Productivity seems more negatively associated with labor share than the technology indices. Plant size is positively related to production labor share.[61]

An arguably more appropriate measure of production labor share is $1 - \alpha_i$, see equation (6). Table 15 shows the results of the previous analysis using this measure as the dependent variable. The results indicate that technology is more strongly associated with this measure of labor share, which can be explained by comparing (6) and (7): if $\gamma < 1$ then $\gamma(1-\alpha_i) < (1-\alpha_i)$, meaning that a given change in α_i should imply a smaller decline in production labor cost's share in revenue than in composite input expenditures. Interestingly, $1 - \alpha_i$ is positively associated with

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

productivity and negatively associated with plant size, which are the opposite of the estimated signs for these variables in the analysis of production labor's revenue share in Table 14. This latter result suggests that the choice of labor share is non-trivial because it may have important consequences in subsequent analyses.

Next, consider the relationship between productivity and automation. The model in Section 4 is agnostic about the relationship between technology, α_i, and TFP, θ_i. To be more accurate, the fully specified model implies no correlation between α_i and θ_i. However, a systematic relationship may be detected between the two variables 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.[62] 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 16 contains the estimated coefficients, which indicate that more productive and larger plants tend to be also more automated.

Putting all the results together, a simple characterization of the relationship between productivity, labor share, and automation emerges. More automated production units tend to be larger, younger, and more productive. Higher automation is associated with lower production labor share, more so if the latter is properly measured. This overall picture suggests that the decline in labor share over time can in part be due to increasing adoption and use of labor-saving technologies by newer, larger, and more productive plants. As a result, the increasing dominance of large and productive businesses (superstar firms) in the economy can be a key driver of the fall in labor share, with relatively higher degree of automation in these businesses contributing further to that decline.

It is important to note that assuming a constant-returns-to-scale Cobb-Douglas production – the most common specification in the literature – implies a different nexus for productivity, automation, and production labor share. Table A4 in Appendix A.2 provides the estimated coefficients from bivariate regressions of labor share and technology measures on productivity. Note that both CES_{EN} and CD_{CRS} are negatively associated with the share of production labor in revenue, but the latter has a much stronger negative association. This is a consequence of the fact that CD_{CRS} underestimates the true underlying productivity, more so for more automated plants, as discussed in the previous section. More importantly, while the share of production labor in composite input expenditure is negatively associated with CES_{EN}, it has no significant association with CD_{CRS}. Finally, automation is positively related to both CES_{EN} and CD_{CRS}, but the relation is stronger in the case of CES_{EN}. These findings indicate that the specification of the production function matters significantly for assessing key relationships between implied variables. One con-

[62] For instance, automation may enhance managerial ability, inventory management, or coordination in factory floor. These are factors not fully captured by standard measures of input usage.

sequence is that heterogenous agent models that aim to capture specific relationships regarding productivity, automation, and capital-labor substitution can be substantially misinformed if an incorrectly specified production function is used to estimate the targeted moments.

7 Conclusion

There is a growing body of theoretical and empirical work on the aggregate effects of automation on manufacturing employment, output, and productivity. However, direct micro-evidence on the connection between automation, labor share and productivity has been limited, due mainly to lack of data sources that contain plant-level information on automation. In particular, little is known about what type of plants rely more on automation, and whether these plants indeed utilize relatively less labor, particularly production workers. This paper provides new evidence on the nexus of automation, labor share, and productivity using plant level data from the U.S. Census Bureau's 1991 Survey of Manufacturing Technology, a unique dataset that contains a rich set of measures on automation-related technology use and investment.

A number of stylized facts from the 1991 SMT indicate that plants with greater use of, and investment in automation have higher capital share and lower production labor share. More automated plants also have a lower fraction of production workers, higher labor productivity and higher wages, and these relationships are more pronounced for production labor than non-production labor. These patterns are consistent with theories that emphasize the replacement of production labor with automation.

Motivated by the stylized facts, a model of plant-level production is presented and estimated. A distinguishing feature of the approach is a CES production function with endogenous technology choice. The key idea behind this modeling device is that plants choose the relative weight of production labor and capital in response to differences in the relative input prices they face. While this is a deviation from the more frequently used assumptions of Cobb-Douglas or CES functions where technology is exogenously given and plants face identical input prices within an industry, it is necessary if the variations in the capital-labor ratio and the relative share of the two inputs are indeed driven by differences in relative input prices. The other distinguishing feature of the analysis is the use of plant-level indicators of automation in the estimation process.

The elasticity of substitution estimates fall in a range where the production labor share is decreasing and the capital-labor ratio is increasing in the relative price of production labor. In addition, the total factor productivity distribution implied by endogenous CES technology differs significantly from those based on more standard Cobb-Douglas and CES specifications. The estimated plant-level productivity is positively correlated with the degree of automation and negatively correlated with production labor's share in revenue. Furthermore, using a common Cobb-Douglas specification results in a very different assessment of the nexus of productivity, production labor share, and automation, indicating that the specification of production technology matters.

The findings also tie into the growing literature on the emergence and evolution of superstar firms. Indeed, the analysis reveals that plants with higher degree of automation tend to experience larger five- and ten-year declines in labor share and bigger surges in labor productivity. Given the positive connection between productivity and growth established in previous studies, the results suggest that more intense adoption and use of automation may be associated with the rise of superstar firms. For the economy as a whole, the diffusion of automation may be one mechanism through which successful businesses with high productivity lower their labor costs, leading to a decline in the aggregate share of labor.

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. The Race between Machine and Man: Implications of Technology for Growth, Factor Shares and Employment. *American Economic Review*, 108(6): 1488–1542, 2018b.

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

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, 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.

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.

Basker, Emek, Lucia Foster, and Shawn Klimek. Customeremployee substitution: Evidence from gasoline stations. *Journal of Economics & Management Strategy*, 26:876–896, 08 2017.

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.

Brown, Byron W. Test from Cobb-Douglas and CES Production Functions. *International Economic Review*, 11(1):77–83, Feb 1970.

Brynjolfsson, Erik and Lorin Hitt. Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending. *Management Science*, 42(4).

Brynjolfsson, Erik and Shinkyu Yang. Information Technology and Productivity: A Review of the Literature. *Advances in Computers*, 43:179–214, 1996.

Brynjolsson, Erik and Lorin Hitt. Computing Productivity: Firm-level Evidence. *Review of Economics and Statistics*, 85(4):793–808, 2003.

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.

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.

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. CEP Discussion Paper 1335, Centre for Economic Performance, LSE, March 2015.

Haltiwanger, John and Zoltan Wolf. Reallocation and Productivity: Structural Estimation and Interpretation. 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.

Kmenta, Jan. On Estimation of the CES Production Function. *International Economic Review*, 8:180–189, June 1967.

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.

Syverson, Chad. What Determines Productivity? *Journal of Economic Literature*, 49(2):326–365, 2011.

Table 1: Summary statistics for 2-digit manufacturing industries, 1991

Industry	SIC Code	Prod. Lab. Share	Non-Prod. Lab. Share	Cap. Share	Cap. Share/ Prod. Lab. Sh.	Avg. TFP (5-factor)
Food & Kindred	20	0.06	0.03	0.32	5.6	1.00
Tobacco	21	0.03	0.02	0.24	7.6	1.15
Textile Mill	22	0.13	0.04	0.48	3.7	1.03
Apparel & Other Textile	23	0.15	0.06	0.17	1.1	0.99
Lumber & Wood	24	0.14	0.05	0.35	2.6	0.97
Furniture & Fixtures	25	0.16	0.08	0.30	1.9	0.95
Paper & Allied Products	26	0.10	0.05	0.70	6.7	0.97
Printing & Publishing	27	0.11	0.14	0.35	3.2	0.94
Chemicals & Allied	28	0.05	0.06	0.52	10.2	1.00
Petroleum and Coal	29	0.02	0.01	0.37	20.8	0.93
Rubber & Misc. Plastics	30	0.13	0.07	0.46	3.5	0.98
Leather	31	0.13	0.06	0.24	1.8	0.97
Stone, Clay & Glass	32	0.15	0.06	0.69	4.7	0.96
Primary Metal	33	0.11	0.05	0.85	7.4	0.96
Fabricated Metal	**34**	**0.15**	**0.08**	**0.51**	**3.4**	**0.94**
Industrial Machinery	**35**	**0.12**	**0.11**	**0.51**	**4.4**	**0.99**
Electronic & Other Electric	**36**	**0.10**	**0.11**	**0.49**	**4.8**	**0.98**
Transportation	**37**	**0.10**	**0.07**	**0.35**	**3.7**	**1.00**
Instruments & Related	**38**	**0.09**	**0.16**	**0.37**	**4.0**	**1.03**
Miscellaneous	39	0.12	0.09	0.31	2.6	1.03
Avg. (All)		0.11	0.07	0.43	5.2	0.99
Avg. (SMT)		0.11	0.11	0.45	4.1	0.99

Source: NBER-CES Manufacturing Productivity Database. Notes: 2-digit industries covered by SMT are boldfaced. Average TFP is calculated across 4-digit industries within each 2-digit industry.

Table 2: Diffusion rates of technologies covered in the Survey of Manufacturing Technology (SMT)

Technology	Diffusion Rate (%) 1988	1993
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
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
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
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 3: Questions on technology use and investment in 1991 Survey of Manufacturing Technology (SMT)

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
	≥ 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
	≥ $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
	≥ 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 (≥ 75%)	4

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

Table 4: Multivariate regressions of labor share on plant characteristics

	Labor Share			Fraction of Workers
	All	Production	Non-production	in 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

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 3. 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 5: Multivariate regressions of labor productivity on 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

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 3. Technology index II is based only on the investment question (Question 2). All regressions include other plant characteristics – see notes to Table 4 for the list. N is rounded for disclosure avoidance.

Table 6: Multivariate regressions of average wage (salaries and wages per employee) on 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 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 3. Technology index II is based only on the investment question (Question 2). All regressions include other plant characteristics – see notes to Table 4 for the list. N is rounded for disclosure avoidance.

Table 7: The relationship between change in production labor share and automation

	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

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 3. Technology index II is based only on the investment question (Question 2). All regressions include other plant characteristics as controls – see notes to Table 4 for the list. N is rounded for disclosure avoidance.

Table 8: The relationship between the change in production labor productivity and automation

	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: 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 3. Technology index II is based only on the investment question (Question 2). All regressions include other plant characteristics – see notes to Table 4 for the list. N is rounded for disclosure avoidance.

Table 9: Estimates of σ, the elasticity of substitution between capital and production labor

	SMT (IV)		ASM (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. 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.

Table 10: 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 (13)			
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 11: Descriptive statistics of productivity measures, SMT 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

Notes: Outliers are filtered in yearly distributions. Industry-year effects are removed. N is rounded for disclosure avoidance.

Table 12: Correlations among productivity distributions, SMT

1991					
	CD_{CRS}	CR_{NCRS}	CES_{FOC}	CES_{EN}	CES_{EX}
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

Notes: Industry-year effects are removed from productivity measures.

Table 13: 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: 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 14: The relationship between productivity and production labor share of revenue

	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

Notes: All continuous variables in logs. Standard error 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 3. Technology index II is based only on the investment question (Question 2). Specifications III and V include other plant characteristics – see notes to Table 4 for the list. Productivity, technology indices and employment are expressed as deviations from industry means. N is rounded for disclosure avoidance.

Table 15: The relationship between productivity and production labor share of the composite input expenditure

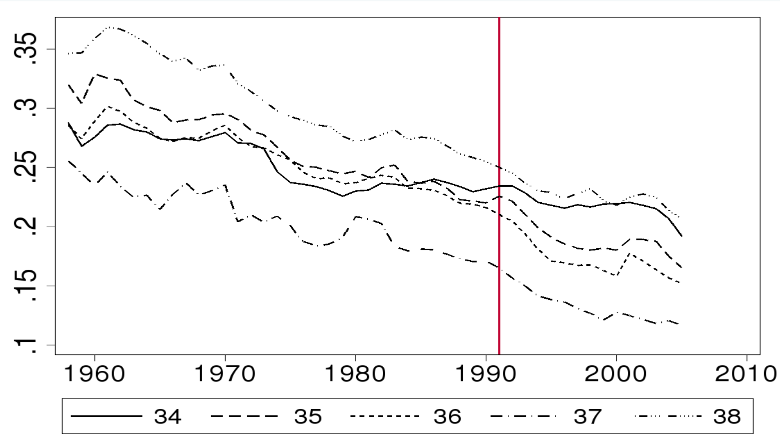
	Production Labor Share of Composite Input Expenditure				
	I	II	III	IV	V
CES_{EN}	0.184***	0.299***	0.298***	0.303***	0.302***
	[0.013]	[0.018]	[0.018]	[0.018]	[0.018]
technology index I		-0.072***	-0.071***		
		[0.014]	[0.014]		
technology index II				-0.127***	-0.126***
				[0.015]	[0.015]
employment		-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

Notes: All continuous variables in logs. Standard error 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 3. Technology index II is based only on the investment question (Question 2). Specifications III and V include other plant characteristics – see notes to Table 4 for the list. Productivity, technology indices and employment are expressed as deviations from industry means. N is rounded for disclosure avoidance.

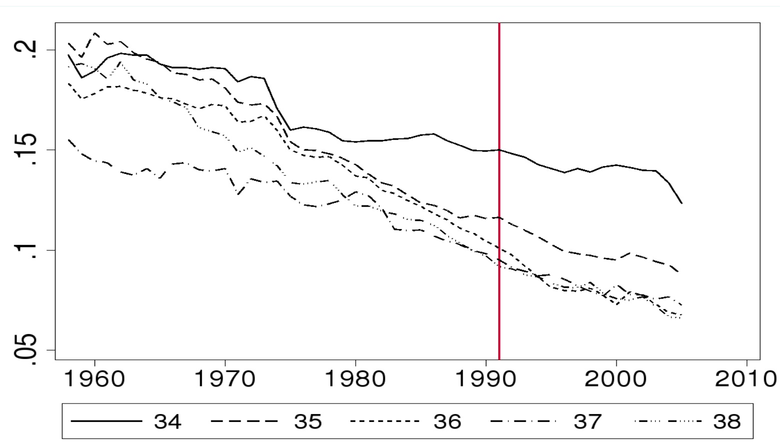
Table 16: The relationship between automation and productivity

	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

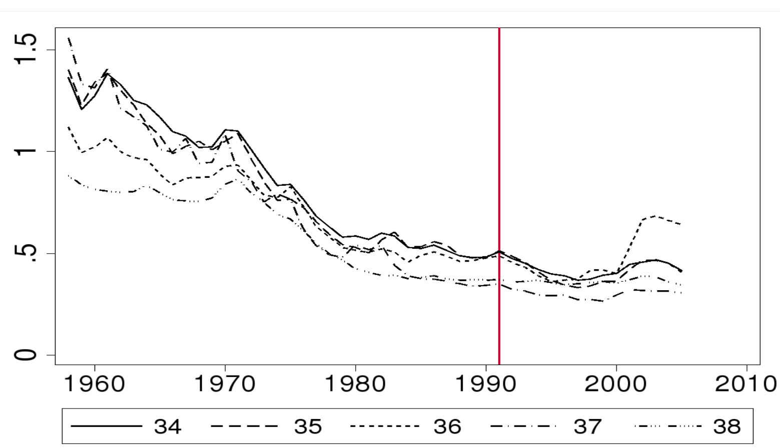
Notes: All continuous variables in logs. Standard error are clustered by 4-digit SIC industry. $(^{*}), (^{**}), (^{***})$ indicate significance at 10, 5, and 1% level, respectively. Technology index I is based on all 4 survey questions in Table 3. Technology index II is based only on the investment question (Question 2). Specifications III includes other plant characteristics – see notes to Table 4 for the list. Productivity, technology indices and employment are expressed as deviations from industry means. N is rounded for disclosure avoidance.



(a) Labor

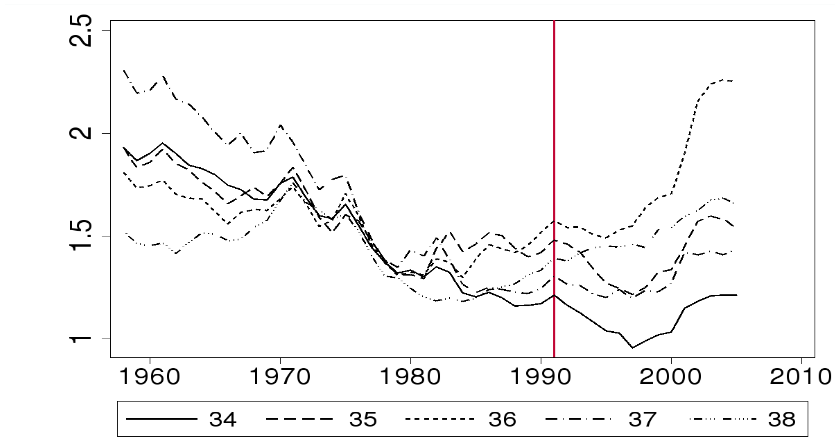


(b) Production labor

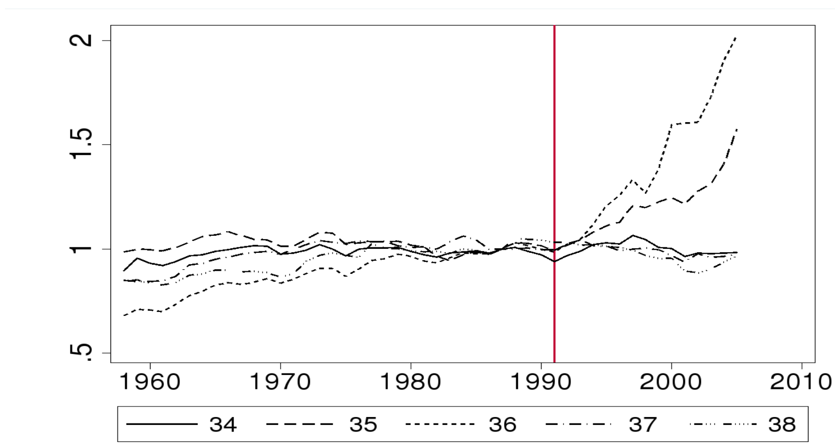


(c) Capital

Figure 1: The evolution of the shares of capital and labor costs in the total value of vhipments, 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



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



(b) Average TFP across 4-digit industries

Figure 2: The evolution of relative capital and labor shares and average TFP, in logs.
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

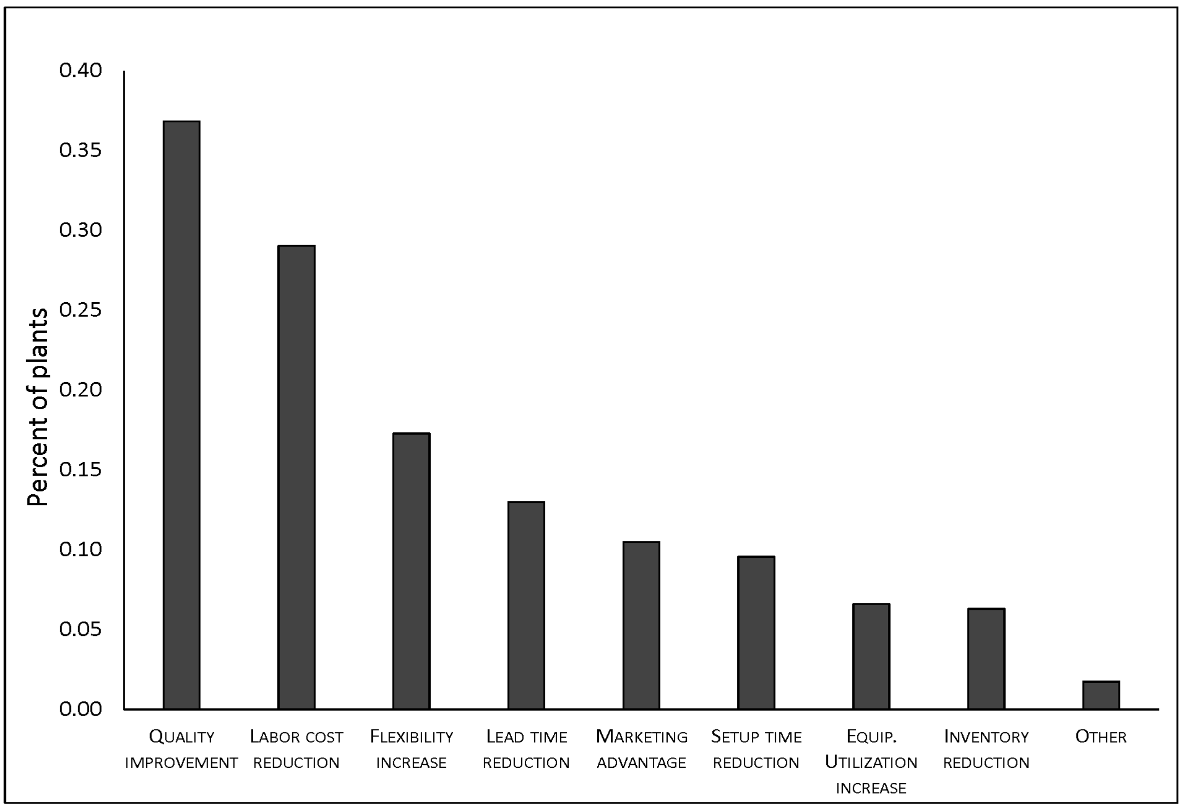
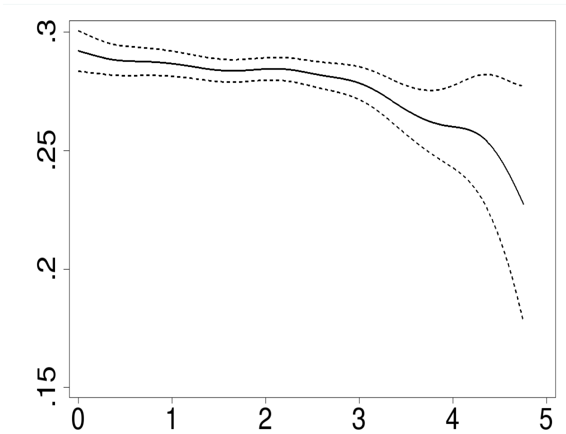
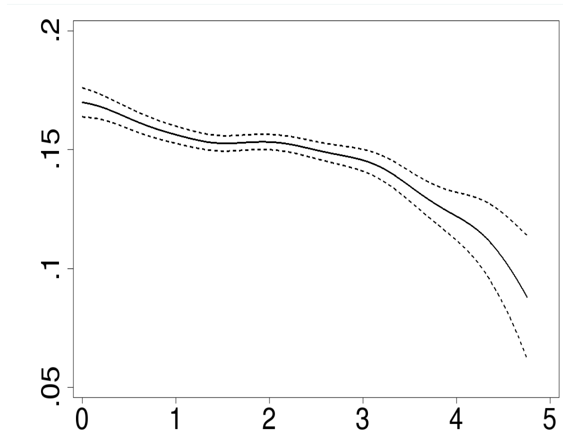


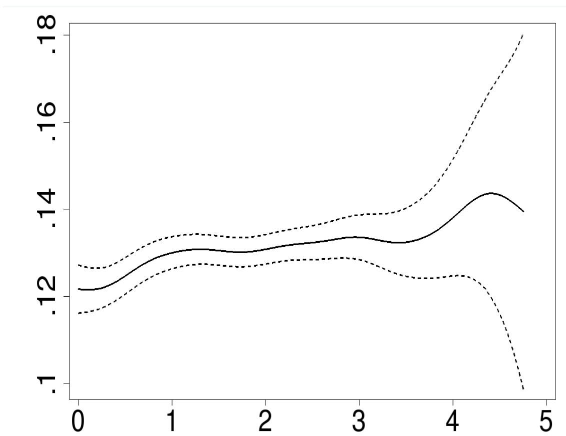
Figure 3: Benefits derived from the use of automation-related technologies – subjective assessment of plants.
Source: Survey of Manufacturing Technology



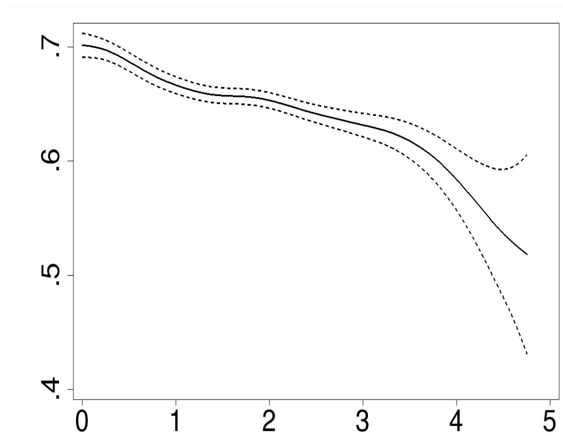
(a) Labor share



(b) Production labor share

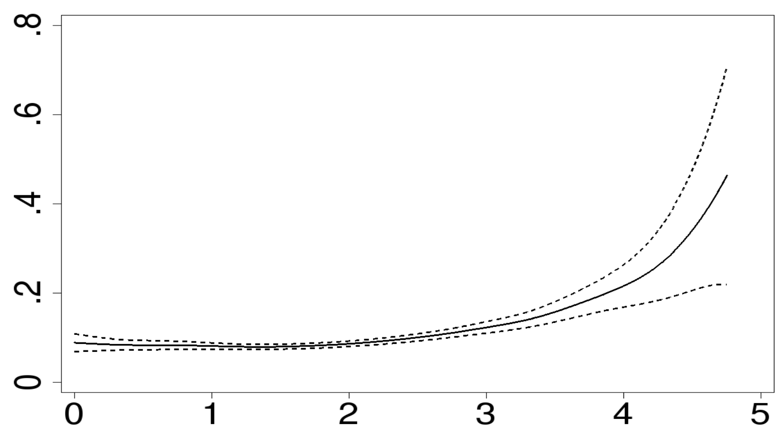


(c) Non-production labor share

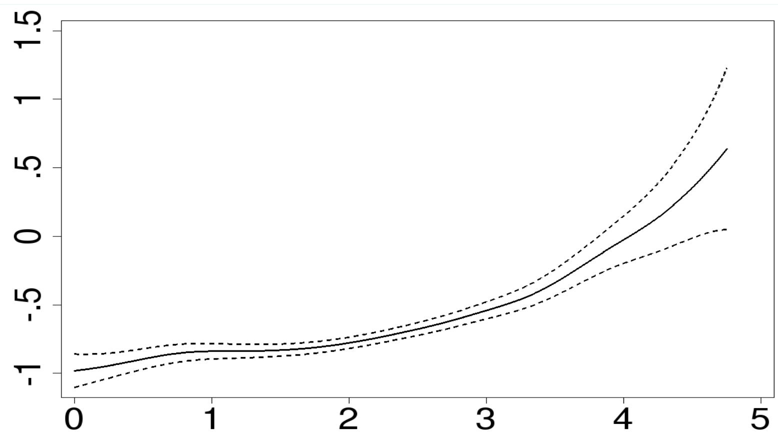


(d) Production worker fraction

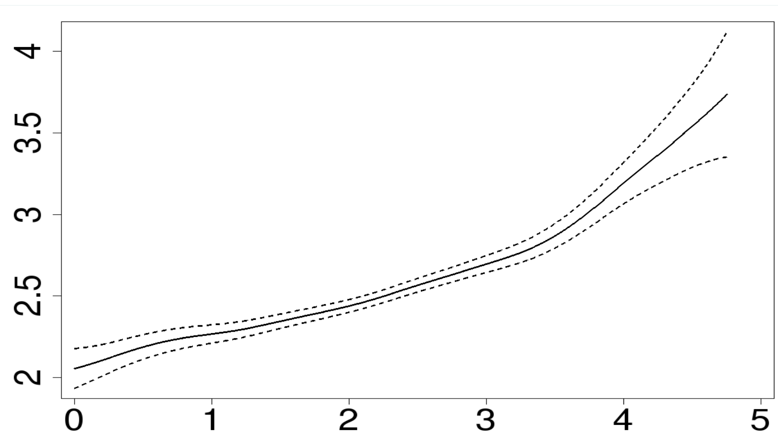
Figure 4: Labor usage as a function of technology index.
Labor share is the share of labor costs in value of shipments. The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Dotted lines show 95% confidence intervals for local polynomial smoothing.
Source: Survey of Manufacturing Technology, Annual Survey of Manufactures, Census of Manufacturing



(a) Capital share

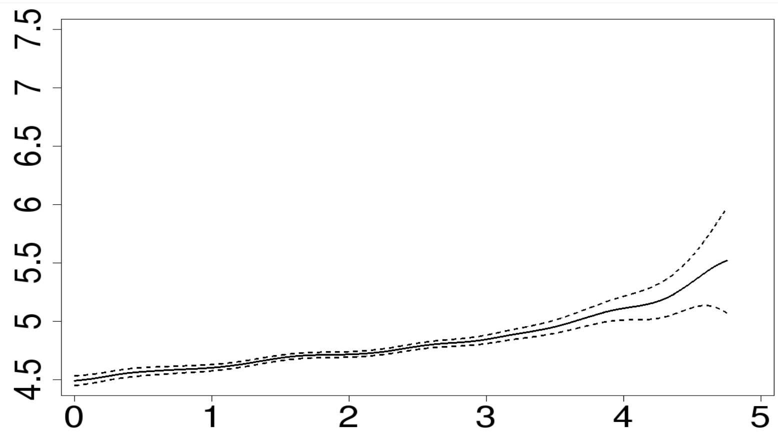


(b) Ratio of capital cost share to production labor share, in logs

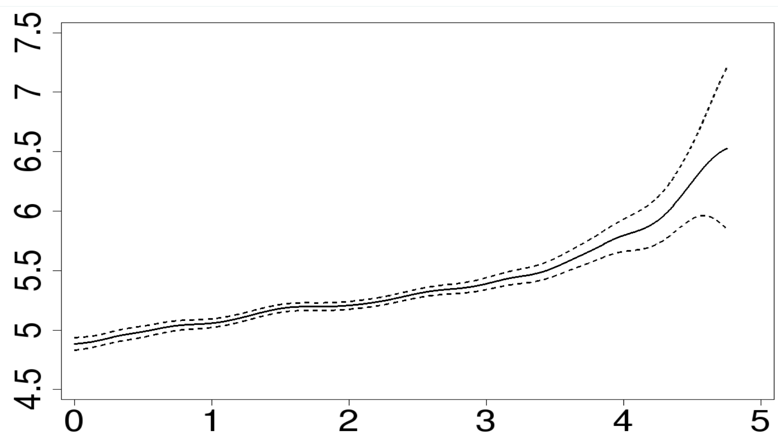


(c) Ratio of capital to production labor, in logs

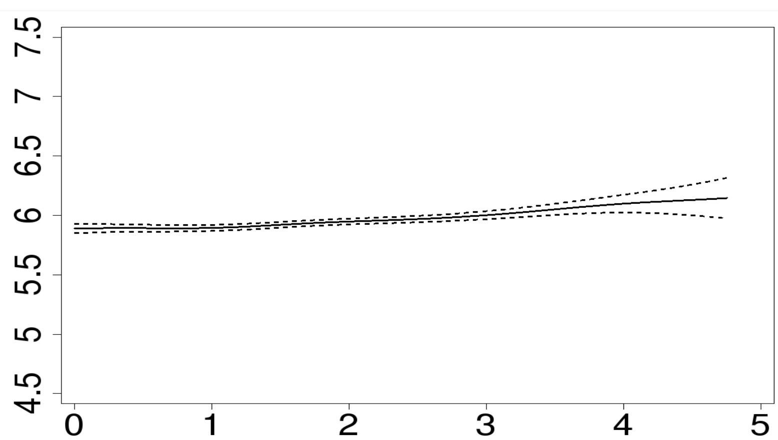
Figure 5: Capital usage as a function of technology index.
Capital share is the share of capital costs in value of shipments. The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Dotted lines show 95% confidence intervals for local polynomial smoothing.
Source: Survey of Manufacturing Technology, Annual Survey of Manufactures, Census of Manufacturing



(a) Labor

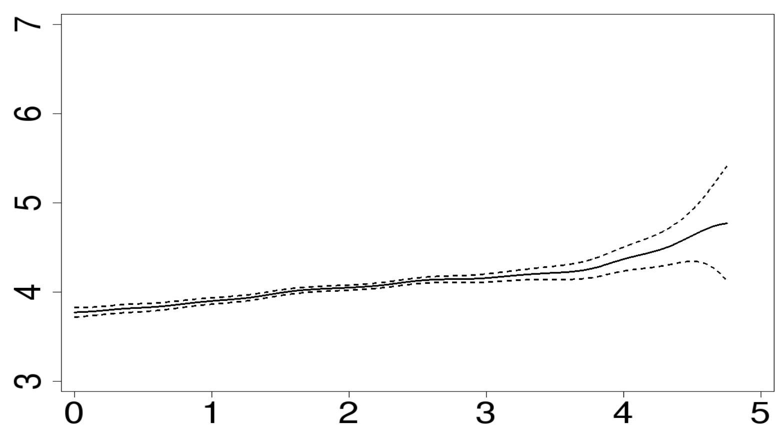


(b) Production labor

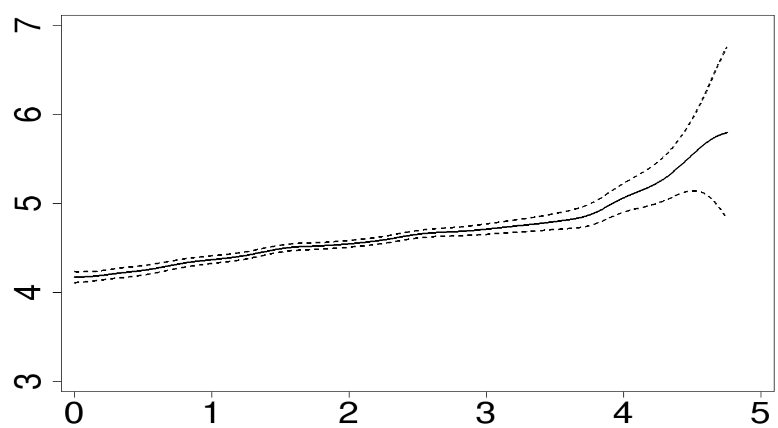


(c) Non-production labor

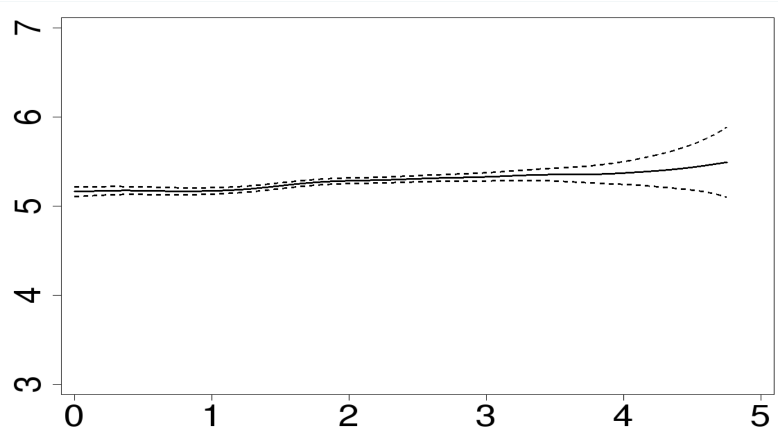
Figure 6: Value of shipments per worker (in logs) as a function of technology index.
The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Revenue is measured as Total Value of Shipments. Dotted lines show 95% confidence intervals for local polynomial smoothing.
Source: Survey of Manufacturing Technology, Annual Survey of Manufactures, Census of Manufacturing



(a) Labor

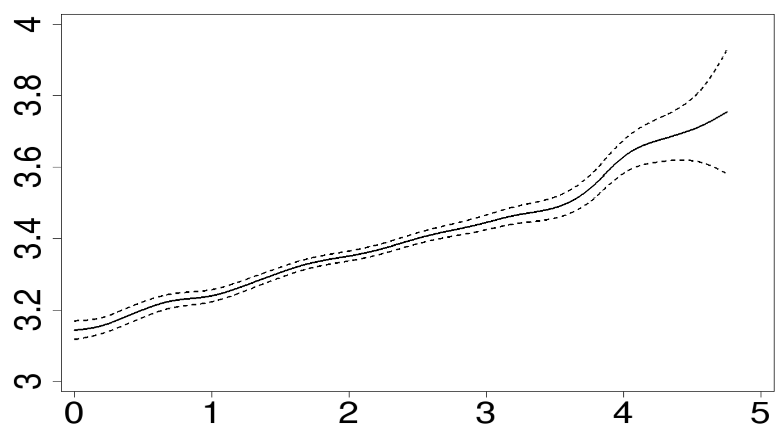


(b) Production labor

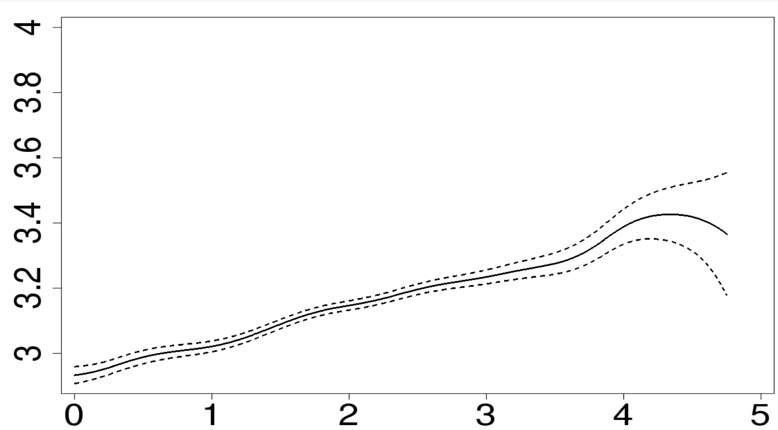


(c) Non-production labor

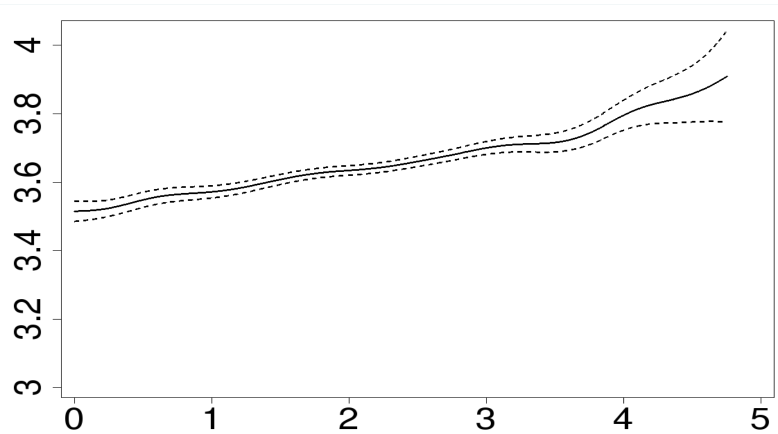
Figure 7: Value added per worker (in logs) as a function of technology index.
The technology index is defined as a plant-specific average of categorical and instance measures of automation technologies, see section 2 for details. Dotted lines show 95% confidence intervals for local polynomial smoothing.
Source: Survey of Manufacturing Technology, Annual Survey of Manufactures, Census of Manufacturing



(a) Labor



(b) Production labor



(c) Non-production labor

Figure 8: 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 Appendix

A.1 Specification error

This appendix studies some properties of the specification error when TFP is estimated using the common model of Cobb-Douglas production function with constant returns to scale

$$Q_i = \theta_i K_i^{\beta} L_{pi}^{1-\beta}, \tag{15}$$

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

$$Q_i = \theta_i[\alpha_i^{2/\sigma} K_i^{\rho} + (1-\alpha_i)^{2/\sigma} L_{pi}^{\rho}]^{\gamma/\rho}. \tag{16}$$

Note that (15) is a typical specification used in the literature on productivity estimation. The two production functions above abstract from the variable inputs L_n, M and E used in (2), mainly for ease of exposition. Including them does not change the main conclusions regarding the theoretical relationship between specification error and technology.[63] For notational ease, the subscript i denoting a plant is omitted for the rest of this appendix.

Let $k = \ln K$, $l_p = \ln L_p$, and $\widehat{\sigma} = 1/(1-\widehat{\rho})$. The difference between the estimated Cobb-Douglas-based log TFP and the estimated CES-based log TFP is given by

$$\begin{aligned}\Delta &= \widehat{tfpr}^{CD} - \widehat{tfpr}^{CES} \\ &= \frac{\widehat{\gamma}}{\widehat{\rho}} \ln[\alpha^{2/\widehat{\sigma}} K^{\widehat{\rho}} + (1-\alpha)^{2/\widehat{\sigma}} L^{\widehat{\rho}}] - \widehat{\beta} k - (1-\widehat{\beta}) l_p.\end{aligned}$$

After some manipulation of terms, one can rewrite Δ as

$$\begin{aligned}\Delta &= \left(\frac{\widehat{\gamma}}{\widehat{\rho}} \ln[\alpha^{2/\widehat{\sigma}} K^{\widehat{\rho}} + (1-\alpha)^{2/\widehat{\sigma}} L^{\widehat{\rho}}] - \frac{\gamma}{\rho} \ln[\alpha^{2/\sigma} K^{\rho} + (1-\alpha)^{2/\sigma} L^{\rho}]\right) \\ &\quad + \left(\frac{\gamma}{\rho} \ln[\alpha^{2/\sigma} K^{\rho} + (1-\alpha)^{2/\sigma} L^{\rho}] - \beta k - (1-\beta) l_p\right) \\ &\quad + \left(\beta k + (1-\beta) l_p - \widehat{\beta} k - (1-\widehat{\beta}) l_p\right) \\ &= \Delta_{CES}^{E} + \Delta^{S} - \Delta_{CD}^{E},\end{aligned}$$

where Δ^S is the specification error due to functional form assumption, and Δ_{CES}^E and Δ_{CD}^E are the estimation (sampling) errors associated with the CES and CD specifications, respectively. The

[63] However, depending on how the elasticities of L_n, M and E are estimated in the case of (15) versus (16), there will be additional discrepancy between the estimated productivities based on CD versus CES specifications.

estimation errors can be written as

$$\Delta^E_{CD} = (\widehat{\beta} - \beta)(k - l_p),$$

and

$$\Delta^E_{CES} = \frac{\widehat{\gamma}}{\widehat{\rho}} \ln[\alpha^{2/\widehat{\sigma}} K^{\widehat{\rho}} + (1-\alpha)^{2/\widehat{\sigma}} L^{\widehat{\rho}}] - \frac{\gamma}{\rho} \ln[\alpha^{2/\sigma} K^{\rho} + (1-\alpha)^{2/\sigma} L^{\rho}].$$

Consider now on the specification error

$$\Delta^S = \frac{\gamma}{\rho} \ln[\alpha^{2/\sigma} K^{\rho} + (1-\alpha)^{2/\sigma} L^{\rho}] - \beta k - (1-\beta) l_p. \tag{17}$$

If the estimation errors Δ^E_{CES} and Δ^E_{CD} are small, the specification error is closely approximated by replacing the true parameters in (17) with their estimates

$$\widehat{\Delta^S} = \frac{\widehat{\gamma}}{\widehat{\rho}} \ln[\alpha^{2/\widehat{\sigma}} K^{\widehat{\rho}} + (1-\alpha)^{2/\widehat{\sigma}} L^{\widehat{\rho}}] - \widehat{\beta} k - (1-\widehat{\beta}) l_p.$$

In (17), a first order Taylor series approximation to the term $\ln[\alpha^{2/\sigma} K^{\rho} + (1-\alpha)^{2/\sigma} L^{\rho}]$ around $\rho = 0$ yields

$$\begin{aligned}
\ln[\alpha^{2/\sigma} K^{\rho} + (1-\alpha)^{2/\sigma} L^{\rho}] &= \frac{d}{d\rho} \ln[\alpha^{2(1-\rho)} K^{\rho} + (1-\alpha)^{2(1-\rho)} L^{\rho}]\Big|_{\rho=0} \rho + \xi, \\
&= \{Bk + (1-B) l_p - 2\left[B \ln\alpha + (1-B)\ln(1-\alpha)\right]\} \rho + \xi,
\end{aligned}$$

where

$$B = \frac{\alpha^2}{\alpha^2 + (1-\alpha)^2}, \tag{18}$$

and ξ is the approximation error for the Taylor series. One can thus approximate the specification error, Δ^S, as follows

$$\widetilde{\Delta^S} = (\gamma B - \beta)(k - l_p) + (\gamma - 1) l_p - 2\gamma\left[B \ln\alpha + (1-B)\ln(1-\alpha)\right]. \tag{19}$$

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 itself. Note also that when $\gamma = 1$, the second term vanishes – assuming CRS in the CES specification (17) implies that $\widetilde{\Delta^S}$ is composed of only the first and third terms. To study the contribution of the third term, let

$$D(\alpha) = -2\left[B \ln\alpha + (1-B)\ln(1-\alpha)\right] > 0.$$

The third term is then $\gamma D(\alpha)$. $D(\alpha)$ is non-monotonic function of α. It achieves its maximum at $\alpha = 0.5$, which is equal to 1.386γ, and its minimum of zero at $\alpha = 0$ or $\alpha = 1$.[64] The exact shape of $D(\alpha)$ is shown in figure A.1.

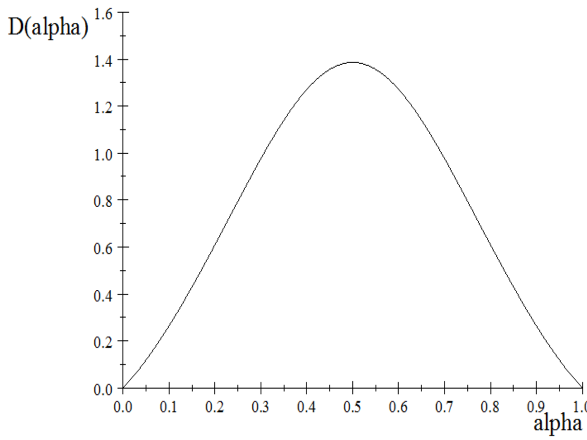


Figure A1: The shape of $D(\alpha)$.

The term $\gamma D(\alpha)$ is bounded from above by 1.386γ, and its contribution to the specification error will be dominated by the first two terms, especially when γ is relatively small.[65] The rest of the appendix assumes that this is the case.

What is the sign of the specification error, Δ^S? Consider the approximation, $\widetilde{\Delta^S}$, in (19). Because $\gamma < 1$, the second term in (19) is negative (for $l_p > 0$, or equivalently, $L_p > 1$). For $(k - l_p) > 0$ (i.e. $(K/L_p) > 1$), the first term in (19) can be positive or negative, depending on whether $\gamma B - \beta$ is positive or negative.[66] If the first term is negative, $\widetilde{\Delta^S}$ is then negative for all plants with at least one production worker and capital-production labor ratio greater than one.[67] In the empirical results discussed in Section 6, the estimated specification error, $\widehat{\Delta^S}$, turns out to

[64]Note that application of L'Hopital rule results in

$$\begin{aligned} \lim_{\alpha \to 0} B \ln \alpha &= \lim_{\alpha \to 0} \frac{\alpha^2 \ln \alpha}{\alpha^2 + (1-\alpha)^2} = 0, \\ \lim_{\alpha \to 1} (1 - B) \ln(1 - \alpha) &= \lim_{\alpha \to 1} \frac{(1-\alpha)^2 \ln(1-\alpha)}{\alpha^2 + (1-\alpha)^2} = 0. \end{aligned}$$

[65]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 $\gamma D(\alpha)$.

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

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

be negative for nearly the entire set of plants for which the two measures were calculated.

Another important question is whether the specification error is exacerbated for plants with higher degree of automation. Consider the first term in $\widetilde{\Delta^S}$. Both $(\gamma B - \beta)$ and $(k - l_p)$ are increasing functions of α, by the definition of B in (18), and by equations (3) and (4) for $\sigma < 1$. Now, let $F(\alpha) = \gamma B - \beta$ and $G(\alpha) = k - l_p$. Note that $F' > 0$ and $G' > 0$ given the preceding discussion. Then, the rate of change of first term with α is given by $F'G + G'F$. When $F > 0$, then $F'G + G'F > 0$. When $F < 0$, $F'G + G'F$ can be positive or negative. As a result, the first term $(\gamma B - \beta)(k - l_p)$ can be an increasing or decreasing function of α. The second term in (19), $(\gamma - 1)l_p$, can also be increasing or decreasing in α, depending on how l_p changes with α. The model has no prediction on the direction of this change. For instance, $(\gamma - 1)l_p$ is decreasing in α if large plants (large l_p) are also the ones with higher α.[68] The overall sign of the change in the specification error as α increases then depends on the behavior of the first and second terms.

Empirical results reveal that the specification error becomes more negative as plant technology (automation) increases. In other words, specification error tends to be larger (in absolute value) for more technologically advanced plants, and the $\mathrm{CD_{CRS}}$ tends to underestimate the underlying TFP (as estimated by $\mathrm{CES_{EN}}$) more for such plants. Table A1 contains the projections of $\Delta = \mathrm{CD_{CRS}}$-$\mathrm{CES_{EN}}$ calculated based on the sample of plants in the analysis on key components of Δ: technology index (a proxy for α), production labor (l_p), capital-production labor ratio ($k - l_p$), 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 $(\gamma B - \beta)(k - l_p)$ in expression (19). The coefficient estimates for the bivariate projections in Table A1 indicates that Δ 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 is positively associated with the interaction of the technology index with the capital-production labor ratio (Specifications V and VI).

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

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: 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 3. Specification VI includes other plant characteristics aside from employment – see notes to Table 4 for the list. N is rounded for disclosure avoidance.

A.2 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 3. Technology index II is based only on the investment question (Question 2). Second-step includes the plant characteristics listed in Table 4. 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 3. Technology index II is based only on the investment question (Question 2). Second-step includes the plant characteristics listed in Table 4. 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 productivity measures, production labor share, and automation

	Estimated coefficient for:	
Dependent variable:	CES_{EN}	CD_{CRS}
production labor share (revenue)	-0.177***	-0.648***
	[0.028]	[0.034]
	0.02	0.11
production labor share (composite input expenditure)	0.184***	0.012
	[0.013]	[0.020]
	0.07	0.0001
technology index I	0.279***	0.076***
	[0.013]	[0.022]
	0.10	0.003
technology index II	0.360***	0.117***
	[0.013]	[0.021]
	0.17	0.008
N	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 3. 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 and productivity measure, 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.