Structural Transformation and the Rise of Information Technology

Giovanni Gallipoli and Christos A. Makridis*

February 15, 2018

PREPARED FOR THE 2017 CARNEGIE ROCHESTER NYU
CONFERENCE ON PUBLIC POLICY and the JOURNAL OF MONETARY ECONOMICS

Abstract

Has the emergence of information technology changed the structure of employment and earnings in the US? We propose a new index of occupation-level IT intensity and document several long-term changes in the occupational landscape over the past decades. Using Census and US KLEMS micro-data, we show that: (i) the bulk of productivity growth after 1950 is concentrated in IT intensive sectors; (ii) the share of workers in IT jobs has expanded significantly, with little or no pause and IT jobs enjoy a large and growing earnings premium, even after controlling for general task requirements (e.g., cognitive, non-routine); and (iii) the rise of the IT intensive employment share is closely associated with declines in the manufacturing employment share. While earnings premia for college-educated and cognitive/non-routine workers have flattened in the aggregate since 2000, we show that they continued growing in IT intensive jobs and that these jobs have played a key role in accounting for the surge of high tech service labor productivity. We also use our IT intensity index to estimate industry-specific elasticities of substitution between IT and non-IT intensive labor, finding values of 1.6 in manufacturing and 1.3 in services. Finally, we revisit a long-standing question about the relationship between technological progress and productivity and provide evidence that occupation-level IT intensity is positively associated with output growth, especially in the services sector.

*Corresponding author: Christos Makridis, Department of Management Science & Engineering, Department of Economics, Huang Engineering Center, 475 Via Ortega, Stanford University, Stanford, CA 94305-4121, cmakridi@stanford.edu (http://stanford.edu/~cmakridi/). Giovanni Gallipoli: Vancouver School of Economics, University of British Columbia, gallipol@mail.ubc.ca. We thank Murilo Esteves de Santi, Elba Gomez Navas, Wesley Ng and Dante Zago for valuable research assistance. We are especially grateful to Narayana Kocherlakota, Ben Pugsley and an anonymous referee for many constructive comments on earlier drafts of the paper. We also thank David Autor, Raj Choudhury, Paul Gaggl, Robert Valletta, and participants at the 2017 Carnegie-Rochester-NYU Conference on Public Policy.

1 Introduction

Understanding large scale sectoral reallocation is essential for identifying the factors of technological change and growth (Baumol, 1967; Kuznets, 1973), such as the decline in the agricultural employment share (Gollin et al., 2002), the rise of the service economy (Buera and Kaboski, 2012; Buera et al., 2015), and the rise in job polarization (Barany and Siegel, 2017). Leading theories of structural transformation focus on either demand-side mechanisms, which introduce heterogeneity in income elasticities (Kongsamut et al., 2001), or supply-side mechanisms, which highlight heterogeneity in sectoral growth rates (Ngai and Pissarides, 2007; Acemoglu and Guerrieri, 2008).[1]

The United States and other developed countries have experienced a large reallocation of workers from manufacturing to the services sector. This well-documented shift (Herrendorf et al., 2014) has coincided quite closely with the rise of technology-intensive investments (see Figure 1).[2] Of course, the fact that aggregate employment shares and measures of technological diffusion are so highly correlated tells us little about causality. To understand the impact that information technology (IT), such as the introduction of personal computers in the 1970s and the spread of the internet in the 1990s, has had on structural transformation, we introduce a new measurement strategy that ranks jobs based on their IT intensity, and use the measure to study how the rise of IT-intensive jobs has potentially accelerated the pace of structural transformation.

[INSERT FIGURE 1 HERE]

In the first part of the paper, we use scores from the O*NET at the five-digit occupation level to measure the prevalence of IT-intensive tasks in different jobs. Drawing primarily on micro-data from the Census Bureau, we document three novel stylized facts. First, value added per hour worked has surged in IT-intensive sectors between 1950 and 2010, particularly within services where it rose almost fivefold.[3] While IT-intensive manufacturing and services have grown rapidly, low IT sectors have displayed more sluggish growth. Second, we show that the share of workers in IT-intensive job has grown, with little or no pause, from roughly 34% in 1970 to 44% in 2015. We also document that over the same period the raw earnings premium in these jobs has also grown

[1]Combining both of these mechanisms has also received attention; see, for example, Buera and Kaboski (2009). Boppart (2014) and Comin et al. (2015) reconcile these mechanisms by modifying preferences.

[2]These results are consistent with anecdotal evidence from, for example, a LinkedIn survey of employers in 2016 that found 22 out of the 25 skills that recruiters are most interested in are technology and information technology related; https://blog.linkedin.com/2016/01/12/the-25-skills-that-can-get-you-hired-in-2016

[3]As we discuss below, we follow the approach suggested by Herrendorf et al. (2013) by focusing on value added through occupational tasks, rather than on gross output.

from 48% to 66%. These findings are consistent with a broader literature about information technology and firm productivity (Bresnahan et al., 2002; Brynjolfsson and Hitt, 2003; Bloom et al., 2012; Tambe and Hitt, 2012). We also contrast our results with recent evidence about the flattening of the college premium (Valletta, 2016) and the slowing demand for cognitive and non-routine ("C/NR") skills (Beaudry et al., 2016). Despite the flattening of returns in the broader cross-section, there continues to be a steady rise in the college and C/NR premia among IT workers. Third, we exploit geographic variation and find that, between 1980 and 2015, a percentage point (pp) rise in the IT employment share is associated with a 0.35pp decline in the manufacturing employment share under our preferred specification. This conditional correlation is identified from decadal within-county variation in employment shares, after controlling for demographic shifts and average wages. The fact that the conditional correlation holds even after controlling for average wages is consistent with models of structural transformation that emphasize productivity growth over income effects (Ngai and Pissarides, 2007).

These facts are especially interesting in light of the ongoing debate on the causes and consequences of structural transformation in the United States. Herrendorf et al. (2015) and Buera et al. (2015) argue that, over the past decades, differences in technological progress across sectors have been key for structural transformation in the United States. Moreover, as pointed out by Duernecker et al. (2017), both labor productivity growth and the expansion in the employment share of services have continued at a fairly strong pace. This is surprising because the services sector has lower output per worker than manufacturing: models that posit wage equalization predict that increases in expenditures on labor-intensive services would eventually be reflected in lower aggregate productivity growth (Baumol, 1967).[4]

Our index allows us to quantify aggregate labor inputs using jobs with different IT intensity. In the second part of the paper, we use these measures to explore how substitutable high and low IT intensity occupations are in production.[5] We perform this analysis separately for the

[4]As Baumol wrote in in his 1967 article, "...if wages and productivity in the progressive sector both go up 2 per cent per year, costs there will not rise at all. On the other hand, if in the non progressive sector productivity is constant, every rise in wages must yield a corresponding addition to costs-a two per cent cumulative rise in wages means that, year in year out, costs must be two per cent above those of the preceding year. Thus, the very progress of the technologically progressive sectors inevitably adds to the costs of the technologically unchanging sectors of the economy, unless somehow the labor markets in these areas can be sealed off and wages held absolutely constant, a most unlikely possibility."

[5]We formally estimate elasticities of substitution between IT and non-IT labor below. For anecdotal evidence, consider the instance where a mere 14 IT-intensive jobs in a manufacturing plant in Austria are enough to produce 500,000 tons of steel a year:
https://www.bloomberg.com/news/articles/2017-06-21/how-just-14-people-make-500-000-tons-of-steel-a-year-

manufacturing and services sectors, re-examining the large sectoral shifts in employment and wages through the lens of a simple production technology with imperfect substitution of IT and non-IT labor inputs. This model delivers estimable relationships linking price changes to quantity changes, which we then use to quantify the elasticity of substitution between IT and non-IT labor inputs.[6] Our preferred estimates, which correct for dynamic reallocation across sectors, indicate that the elasticity of substitution between IT and non-IT intensive labor inputs is 1.6 in manufacturing and 1.3 in services. We obtain quantitatively similar results when we use the standard estimation approach outlined by, for example, Autor et al. (2008)

Our work also connects with two broad debates about the impact of technology on labor markets and productivity. The first debate examines the decline in productivity growth over the past fifteen years. While some prominent researchers have argued that productivity might simply be mismeasured due to the difficulty of quantifying output in technology-intensive activities (Brynjolfsson and McAfee, 2011; Mokyr, 2014; Bryne et al., 2013; Feldstein, 2015; Hatzius and Dawsey, 2015; Bryne et al., 2016), Syverson (2017) suggests that mismeasurement is an unlikely explanation and Brynjolfsson et al. (2017) argue that new technologies (e.g., artificial intelligence) might just have delayed effects on the real economy since it takes time for the benefits to diffuse.[7]

In the last part of the paper, we provide direct evidence that IT-related productivity effects are present in sectoral real output and value added. Our empirical approach follows that of Acemoglu et al. (2014), who examine productivity growth across manufacturing industries and find that it was surprisingly low in IT-using *industries*. Using our occupation-level IT intensity measure, rather than measures of IT capital, we replicate the analysis of of Acemoglu et al. (2014) for manufacturing and extend it to the services sector. Our estimates suggest that occupation-level IT-intensity is positively associated with output growth, especially in services. This finding points towards an important dimension of heterogeneity—namely, that significant productivity growth may have recently come from IT-intensive *jobs* in the services sector.

The second debate we contribute to highlights growing concerns about automation and income inequality. For example, Harrigan et al. (2016) use administrative data from France to show that

in-austria.

[6]While traditional SBTC contributions have focused on college attainment as a proxy for skill (Katz and Murphy, 1992; Autor et al., 2006) and/or heterogeneity between non-routine & cognitively intensive occupations (Autor et al., 2008; Autor and Dorn, 2013), an important feature of the SBTC intellectual foundation is that it can accommodate alternative, relevant sources of heterogeneity, such as age (Card and Lemieux, 2001) and gender (Acemoglu et al., 2004). See Acemoglu and Autor (2011) for a survey.

[7]See an interesting discussion of related issues here: https://www.wsj.com/articles/silicon-valley-doesnt-believe-u-s-productivity-is-down-1437100700

information technology and technological change have played a major role in facilitating job polarization.[8] However, they focus specifically on technical managers, engineers and technicians (the "techies"), rather than information technology workers as defined by occupational task descriptions, to study polarization in the distribution of jobs. In a similar vein, while some prominent experts have argued that a whopping 47% of U.S. employment is at risk of being displaced by automation (Frey and Osborne, 2013), Acemoglu and Restrepo (2017b) show that continuing automation is consistent with balanced growth, creating new tasks and raising the demand for heterogeneous skills. Even in the presence of short-run displacement of jobs (Acemoglu and Restrepo, 2017a), which clearly has welfare and distributional implications (Eden and Gaggl, 2017), what matters most for structural transformation is the long-run demand for specific tasks and skills. In this sense, our paper contributes towards a better understanding of how information technology shapes the sorting of workers across occupations and the associated returns.

Our findings are also consistent with key findings in Duernecker et al. (2017). These authors suggest that, while Baumol's disease may have somewhat slowed aggregate GDP growth, its effects are incrementally smaller and will be limited in the future because most of the growth will continue to come from sustained demand for the output of high productivity services sectors. Our focus on the growing importance of IT jobs highlights a specific supply-side mechanism that drives higher productivity growth in the services sector. Identifying these occupational shifts will be important in further developing the current models (e.g., Buera and Kaboski (2012), Herrendorf et al. (2013), Boppart (2014), and Comin et al. (2015)) and characterizing the next stage of structural transformation—the move towards information services.

The structure of the paper is as follows. Section 2 introduces the data and measurement strategy. Section 3 employs our IT intensity measure to document three stylized facts about the rise of technology intensive jobs. Section 4 introduces a production-based framework to estimate the elasticity of substitution between IT intensive and non-IT intensive jobs. Section 5 uses our IT intensity measure to assess the role of occupation-level technological change for productivity, revisiting questions related to the Solow paradox. Section 6 concludes.

[8] Their contribution relates to findings in Autor et al. (2003) and Autor and Dorn (2013) in that information and communications technology largely complement cognitive and non-routine skills. Below we also provide evidence suggesting that this may be true, as we show that the earnings premium for college and cognitive/non-routine skills after 2000 has been increasing for information technology workers (but not for non-IT workers).

2 Data and Measurement

2.1 Sources

Occupational Tasks: Measuring IT Intensity.—We draw primarily from O*NET, which is the new companion to the well-known Dictionary of Occupational Titles (DOT) used in prior work (Autor et al., 2003), as our source of data on occupational tasks, skills, and work environment characteristics. O*NET is a survey that the U.S. Department of Labor administers to a random sample of U.S. workers within detailed occupations. Respondents answer questions on an ordinal scale that measures both the importance of, for example, a task, and the frequency at which different tasks occur on the job. Following prior work by Autor et al. (2013b), we take the product of the importance and frequency weights (when available) to generate an overall intensity for each sub-index task. To match these indices to micro-data in our baseline sample, we use the Occupational Employment Statistics (OES) national time series and construct employment-weighted intensities for tasks, skills, and knowledge at the five digit occupation level.

Individual Micro-data: Employment and Wages.—Our primary results draw from the Census Bureau's Decennial Census between 1970 and 2000 and the American Community Survey (ACS) between 2005-2007 ("2005") and 2013-2015 ("2015") accessed through the Integrated Public Use Microdata (IPUMS) data portal at the University of Minnesota. Occupations are measured at a five-digit level of aggregation, which we match with our IT index from O*NET. We also draw on the annual Current Population Survey (CPS) for several exercises that allow us to have higher frequency variation in the number of IT and non-IT workers. In both cases, we mitigate concerns about partial attachment to the labor market by restricting our samples to full-time workers between age 20 and 65, with over $5,000 in annual labor income, at least 20 weeks worked per year, and over $2 real hourly wages. We deflate nominal variables using the 2010 real personal consumption expenditure index.

Industry Classifications.—We define the services sector based on the following industries: utilities, transportation and warehousing, wholesale and retail trade, information, finance and insurance, real estate, professional / scientific / technical services, management of companies and enterprises, administrative and support / waste and remediation services, educational services, healthcare and social assistance, arts / entertainment / recreation, accommodation and food services, and other personal services.[9] We define the manufacturing sector based on NAICS 31-33

[9] https://www.census.gov/econ/services.html

codes (non-durable and durable manufacturing) and include agriculture and mining.[10]

2.2 Measuring Occupation-Level IT Intensity

To gauge the intensity of IT tasks at the occupation level (aggregated at the five-digit level in our baseline), we draw on indices from the O*NET skill, task and knowledge measures, including: knowledge about computers and electronics, activities interacting with computers, programming, systems evaluation skills, quality control analysis, operations analysis, activities with updating and using relevant knowledge, technology design, activities analyzing data and information, activities processing information, knowledge with engineering and technology, activities managing material resources (see Appendix A.1 for a description of each category).

After computing the sum of these sub-indices between 2004 and 2016 (see Appendix A.3.2 for details about time variation of these indices), we construct an average and time-invariant occupation intensity score and standardize it (so that the score has mean zero and standard deviation of unity).[11] We then classify occupations as high IT if they have a standardized score above zero. We harmonize Census occupation classifications from 1970 to 2015 to the five-digit SOC-level such that 8,124,303 of 19,266,737 (42.17%) individuals are classified as working in a high IT intensity job.[12] Figure 2 plots the relationship between occupation-level earnings and standardized IT intensity scores, weighted by employment. The data suggests that a standard deviation increment in IT intensity is associated with a 0.83% increase in annual earnings. Appendix A.3.1 compares and contrasts our measure of IT intensity with alternative proxies of workers' skills, controlling for measures of cognitive / non-routine and social skills used by Autor et al. (2003) and Deming (2017), respectively.

[INSERT FIGURE 2 HERE]

[10]Our results are robust to including the education and healthcare sectors as low skilled in the services sectors and/or excluding agriculture and mining from the manufacturing sector. These latter sub-sectors make little difference because employment shares are quite small in agriculture and mining. Similarly, while education and healthcare clearly contain groups of highly skilled workers, many are not. Our classification on high and low tech services sectors is influenced by the Census (https://www.bls.gov/opub/mlr/2005/07/art6full.pdf), but we also have examined college attainment across these sectors.

[11]Hagerty and Land (2007) find that using equal weights over subset variables for these types of indices provides the greatest robustness and accuracy. The issue of potentially different weights is relatively minor when all are measures are positively correlated.

[12]We take advantage of the fact that IPUMS makes "OCC2010" available, which we manually cross-walk into SOC2010 codes at the five-digit level. Appendix A.1 provides a comparison of the earnings and employment premia under alternative IT intensity definitions based on jobs being either above the mean (baseline) or above the median of the standardized score. Despite some differences, the correlations between the earnings and employment premia under the two definitions are 0.83 and 0.90, respectively.

A key assumption is that our classification of occupations, based on average IT intensity between 2004 and 2016, is time-invariant. This means that we can extrapolate our scores of occupational IT intensity as far back as 1970.[13] While we cannot directly test the validity of this assumption, we gauge its robustness by checking whether occupational IT intensity scores in 2004-2005 have explanatory power for scores in 2015-2016. We find significant persistence with a coefficient of 1.069 (p-value $= 0.00$) and a R-squared of 0.79. The close match between occupational IT intensity in the early 2000s and a decade later does not guarantee that the same would hold over the 1980s and 1990s, but the very strong correlation indicates that relative occupational IT task intensities appear not to change dramatically over time.

2.3 Descriptive Statistics

In Table 1, we document stark gender and college attainment differences between workers in high and low IT-intensity jobs. For example, in 1970 the shares of males and college degree workers in IT-intensive jobs were, respectively, 76% and 33%. This can be contrasted to a share of only 59% males and 4% college graduates in jobs with lower IT intensity. However, these differences have narrowed: for 2010-2015 the shares of males and college grads were 54% and 58% in IT-intensive jobs, versus 52% and 18% in non-IT-intensive jobs.

We also see a large increase in the IT intensive hourly wage premium, growing from \$6.68/hour in 1980 to \$12.55/hour in 2013 (an 87% increase). Controlling for college attainment reduces the premium to \$5.05/hour in 1980 and \$9.11/hour in 2010-2015 (a 65% increase), consistent with the complementarity between skills and computer technologies (Autor et al., 1998). However, the hourly wage premium still remains large. Interestingly, the increase in the hourly wage premium occurred concurrently with, and in spite of, an increase in the hours-worked IT premium. For example, high IT job workers allocated 8.3% more time to market activity in 1980 than workers in low IT jobs, but this gap grew to 11% by 2010-2015.[14] Over this same time period, the IT earnings premium grew from 46.4% to 66.4%. The fact that the earnings premium is rising faster than the hours premium for IT workers suggests that demand is outpacing supply. Appendix A.2 examines how the distribution of employment, hourly wages, and inequality varies across high and low IT intensity jobs. While the distribution of the hourly wage for IT jobs is shifted significantly

[13]For an interesting discussion of related issues see (Atalay et al., 2017).

[14]We have also cross-walked our IT index into the American Time Use Survey (ATUS) and found that, using their more reliable diary-based time use reporting method, IT workers allocate roughly 21 more minutes per day to work activities, which totals nearly 128 more hours worked per year.

to the right, there is interestingly little statistical difference in within-group inequality between high IT and low IT jobs.

[INSERT TABLE 1 HERE]

3 Structural Change and the Emergence of Technology and Information Services

3.1 Heterogeneity in Industry Labor Productivity

How has productivity evolved over time in the manufacturing and services sectors? Although the manufacturing sector exhibits greater total factor productivity (Herrendorf et al., 2015), we document the presence of significant within-industry heterogeneity in labor productivity between high and low IT intensity sectors. We measure sectoral labor productivity as logged value added per hour worked.

We draw on two datasets to examine the evolution of value added. First, we use the U.S. KLEMS 2013 release to obtain sectoral total hours worked and value added from 1950 to 2010, which we deflate using an annual GDP deflator (normalized to 2009). Second, we use our Census micro-data samples for 2000, 2006-2008, and 2013-2015 to produce a weighted z-score of IT intensity at a three-digit NAICS industry level, classifying sectors with an average z-score above zero as high IT intensity; see Appendix A.3.3 for further details on the data construction and list of sectors used in the classifications. Figure 5 displays the results of our analysis. First, labor productivity in high IT intensity manufacturing and services are both higher than labor productivity in low IT sectors. In particular, the IT services sector's productivity appears to catch up with high tech manufacturing after 1980. Second, the long-term increase in value added per hour is highest in the IT intensive services, which grow from a value of $9.43 per hour worked in 1950 to over $57 in 2010 (a fivefold increase). In contrast, IT intensive manufacturing's value added grows from $19.4 per hour worked in 1950 to $83 in 2010 (a threefold increase).

3.2 Heterogeneity in Employment and Earnings across Occupations

Having motivated our emphasis on high-tech services, we next focus on occupation heterogeneity within both the manufacturing and services sectors. Using micro-data between 1970 and 2015 from the Census Bureau, we begin our analysis by plotting the earnings and employment premia between high and low IT-intensive jobs. Figure 6 documents the evolution of the earnings premium for IT-intensive jobs in both manufacturing and services sector. Whereas the earnings premium grew from 50.4% to 63.6% in the manufacturing sector between 1970 and 2013 (a 26.2% rise), it grew from 48.3% to 67.2% in the services sector (a 40% rise). We find very similar patterns when working with hourly wages and residual hourly wages, which purges variation in demographic characteristics and educational attainment.

Turning towards differences in employment, Figure 6 also documents the employment premium for IT-intensive jobs. While manufacturing had only a few IT-intensive jobs in the 1970s, with an employment premium of -0.84, the share of such jobs grew dramatically over time to an employment premium of -0.23 (a 70% rise) by the end of our sample period. Over the same period, the economy also experienced a significant growth in the share of IT-intensive jobs in the services sector, from an employment premium of -0.53 to -0.18 (a rise of 64.6%). These results suggest that, in the face of the well-known decline in the manufacturing employment share (Herrendorf et al., 2014), any job growth (or retention) in the manufacturing sector must have been primarily in IT-intensive jobs. Figure 20 in Appendix A.3.4 presents evidence consistent with this corollary: using the Current Population Survey (CPS), we show that the employment share of IT jobs in the manufacturing sector remained quite stable, despite the rapid decline of the overall manufacturing employment share.

[INSERT FIGURE 20 HERE]

Appendix A.3.4 replicates these earnings premia using a combination of the CPS and OES datasets to illustrate that these premia are comparable in other standard sources. While there are some differences in levels since the level of aggregation differs (three-digit occupation in CPS and six-digit in OES), the qualitative patterns are the same. We also present robustness exercises using a different classification of IT-intensive jobs, namely a more restrictive scheme introduced by the Census Bureau that focuses heavily on science and engineering jobs (Hecker, 2005). We have also estimated these premia controlling for demographic characteristics, such as age and education, and found that our patterns still stand strong. We also find similar results using hourly wages, but defer to annual earnings since hours worked are not reported in 1970.

We now examine how these patterns relate with recent results from Beaudry et al. (2014; 2016) about the decline in demand for non-routine and cognitive (“C/NR”) skills and from Valletta

(2016) about the flattening of the college premium.[15] In particular, has the flattening taken place for all workers, or is the flattening driven by non-IT jobs? After replicating their aggregate results using the Census, we find evidence of the latter. Beginning with educational attainment, Panel A in Figure 7 plots the overall college premium (red) with the college premium for IT workers (blue). Importantly, while all are larger and increasing, the premium is increasing especially for IT workers. For example, between 1970 and 1990, the college premium grew by 9% overall, whereas it grew by 16% for IT workers. Between 1990 and 2015, the college premium grew by 19% overall, but it grew by a remarkable 26% for IT workers. During these years, the supply of IT workers also rapidly increased, suggesting that the supply of labor was not keeping up with the increasing demand for IT skills in the labor market (see Appendix A.3.5 and Figure 25 for these employment shares).

Turning towards C/NR skills, Panel B in Figure 7 plots the overall C/NR premium (red) with the C/NR premium for IT workers (blue). Unlike the college premium, there is a closer association between the overall C/NR and IT-C/NR premia since the majority of occupations considered high C/NR are also high IT. Despite similarity between the two premia in 1970 bordering around 40%, the premia began diverging in the 1990s. In particular, the IT premium among C/NR jobs begins outpacing the overall premium by over 5%. For example, while the overall C/NR premium grew by 14% between 1970 and 1990, it grew by 18% among IT workers. Moreover, the overall C/NR grew by 40% between 1990 and 2015, whereas it grew by 45% among IT workers.

[INSERT FIGURE 7 HERE]

While these plots may appear inconsistent with the evidence in Beaudry et al. (2016) and Valletta (2016) on the declining return to C/NR and college skills post-2000, they are not—they merely highlight an important source of heterogeneity. Turning away from the longer-run phenomena captured in Figure 7, we use annual data from the American Community Survey between 2005 and 2016 to study the returns to IT jobs among the subset of college degree workers and high C/NR jobs. Using these data, we estimate regressions of logged hourly wages on an indicator for high IT intensity, college attainment (and, separately, high C/NR skills), and their interaction for each year, conditional on controls. Figure 8 plots the estimated interaction between (i) IT and

[15] Beaudry et al. (2016) argue that, because of the housing boom in the early 2000s, the rise of employment in the construction and other complementary sectors masked these downward employment trends (Charles et al., 2016). It was not until the Great Recession that these labor market changes became apparent. Autor (2017) points out, however, that the decline in the earnings premium actually began in the 1990s, which is even more of a puzzle since IT investment was still large at the time.

college attainment and (ii) IT and C/NR skills. it is apparent that the hourly wage premium for college workers who also work in IT intensive jobs has grown from 4% to 6.5% between 2005 and 2016. Similarly, the hourly wage premium for workers in IT intensive jobs requiring C/NR tasks has grown from -2% to 2%. While our estimates exhibit large confidence intervals, the coefficients point towards a rise in the premium associated with college and C/NR skills for high IT workers. In this sense, although the average return to college and C/NR skills might be declining overall as Beaudry et al. (2016) and Valletta (2016) point out, there is no evidence that this occurs in the subset of IT intensive jobs.

[INSERT FIGURE 8 HERE]

3.3 Comparison to Alternative Measures

Large differences in earnings, employment and value added across sectors become apparent when we classify occupations using our IT intensity measure. We validate the robustness of these findings in several ways. As mentioned in Section 2, our IT intensity measure defines an occupation as high IT intensity if its standardized z-score index is above zero (the mean value). We have also experimented with other ways of clustering occupations into high and low IT intensity groups, using both Ward's algorithm and K-means/medians. Treating total IT intensity as the outcome variable, these clustering methods did not produce classifications that were significantly different from our baseline approach. Hence, we use our mean z-score approach for its transparency and stability. We define high IT based on values higher than the median z-score. Our results are robust to this alternative definition (see Appendix A.1.2).

Next, we benchmark our results using an industry-level measure of technology prevalence introduced by Hecker (2005). This measure is based on the proportion of scientists, engineers, and technicians at an industry, rather than occupation, level. Hence, Hecker (2005) focuses on a sub-sample of science, technology, engineering, and math (STEM) workers who account for 7% of the jobs in the sample, whereas our IT intensity measure captures a broader set of technology jobs, accounting for 42% in the sample. To contrast these measures, we consider regressions of logged annual earnings and hours worked on indicators of, respectively, high IT intensity or Hecker's high tech industry, controlling for individual covariates.

Table 2 documents these results for the 2000, 2005-07, and 2013-15 sample years.[16] These raw

[16]Unfortunately, we cannot include earlier years since the NAICS codes are not available for them. While we

comparison suggests that the earnings premium in IT intensive jobs is 63%, whereas workers in high tech jobs have 55% higher earnings relative to their counterparts. Once we add controls, the conditional correlations imply premia of 43% and 34%, respectively. We find smaller differences between annual hours worked, amounting to 8% for our IT intensity measure versus 5% for the high tech-based measure, conditional on controls. One limitation of our comparison is that the public Census micro-data does not contain enough information to cover all of the sectors that Hecker (2005) classifies as high tech. Nonetheless these conditional correlations illustrate that these measures capture similar, but not identical, features of earnings and hours premia in the data. Table 7 in Appendix A.3.6 also presents a comparison with an additional and less expansive measure of information technology prevalence within jobs introduced by Beckhusen (2016).

[INSERT TABLE 2 HERE]

We also draw on the Bureau of Economic Analysis (BEA) measures of IT capital expenditures from 2000 to 2015 and examine their correlation with the IT employment share. This check is motivated by evidence of complementarity between high-skilled workers and technology (Griliches, 1969; Autor et al., 1998). Since IT capital expenditures are measured at an industry-level, whereas IT jobs are measured at an occupation-level, we compute the share of IT workers within each three-digit industry-level weighted by the number of individuals observed in each cell. Consistent with prior evidence about the complementarity between skilled workers and capital, Figure 3 shows that the two have a 0.46 correlation and that a one percentage point rise in the share of IT workers at the industry-level is associated with a 2.89% rise in capital expenditures.

[INSERT FIGURE 3 HERE]

While our measure of IT task intensity is clearly detecting meaningful differences across occupations, one concern is that it may be simply correlated with other occupation-specific characteristics, such as non-routine and cognitive features as defined in Autor and Dorn (2013). To verify the extent of this overlap we use more detailed six-digit occupational data from the Occupation Employment Statistics (OES) between 2000 and 2015 and examine differences in the intensity of various other tasks. Figure 4 plots the distribution of intensity for five different tasks in high and low IT intensity occupations. The differences between IT and non-IT are significant for cognitive, technical, social, and general tasks, but not for manual and service task intensities. Even

can cross-walk earlier industry codes to three-digit NAICS classifications, we cannot cross-walk them accurately to four-digit NAICS classifications, which is the basis of the high tech definition in Hecker (2005).

after controlling for differences in task intensities, we still find that IT intensity is significantly associated with earnings. These exercises are documented in Appendix A.3.1. We also present descriptives in Appendix A.3.2 on the time series of IT intensity across jobs.

[INSERT FIGURE 4 HERE]

3.4 IT-Intensive Jobs and Structural Change: Some Evidence

Motivated by these results about the large shifts occurring in the labor market following the emergence of IT-intensive jobs, we next provide some new evidence about their association with structural transformation. In particular, we ask whether increases in the employment share of IT have contributed to the hollowing out of manufacturing employment, after accounting for a variety of confounding factors. Using our decadal Census micro-data to produce county weighted averages between 1980 and 2015, we restrict the sample to counties with over 2,000 survey respondents and run regressions of the form:

$$m_{ct} = \beta X_{ct} + \gamma IT_{ct} + \psi_c + \lambda_t + \epsilon_{ct} \tag{1}$$

where subscripts c and t denote county and decade, m denotes the manufacturing employment share, X denotes a vector of local controlling covariates, IT denotes the IT employment share, and ψ and λ are county and decade fixed effects. A possible endogeneity problem in estimating Equation 1 arises from the fact that higher productivity workers might sort into IT jobs (and thus locations with more of these jobs). To the extent that IT jobs are comparatively less likely to be concentrated in the manufacturing sector, γ may be biased downwards.

While we do not take a strong causal stance on γ, we provide evidence that this relationship is not spurious or entirely contaminated by endogeneity problems. In particular, we begin by introducing several controls that help reduce concerns about omitted variables bias, including: the logged hourly wage, the share of college degree workers, the share of whites, and the share of males. The inclusion of both wages and college shares helps ensure that we are not comparing areas with systematically different labor markets. The share of males addresses the concern of heterogeneous entry among females into different labor markets. The inclusion of county and year fixed effects removes all time-invariant differences across these labor markets, such as exposure to trade shocks or local human capital networks. To address the potential endogeneity arising from reverse causality and unobserved time-varying shocks, we also instrument the contemporaneous

IT employment share with its 1980 value, which is pre-determined for $t > 1980$.

Table 3 documents our results. Column 1 presents the simple unconditional correlation, which suggests that a percentage point rise in the IT share is associated with a 0.67 percentage point decline in the manufacturing share. We subsequently introduce several controls in column 2, lowering the point estimate on the IT share to -0.226. Here, we find that the college share is negatively correlated with the manufacturing share, which is consistent with existing evidence on growing polarization across industries and the skill premium (Autor and Dorn, 2013). However, while these controls help address contemporaneous omitted variables, any time-invariant source of spatial heterogeneity may create attenuation bias. Column 3 introduces county and year fixed effects without controls, which suggests that a percentage point rise in the IT share is associated with a 0.38pp decline in the manufacturing employment share. Recognizing the presence of time-varying unobservables, we once again add our controls, on top of county and year fixed effects, and find that the gradient only declines marginally in magnitude to an associated 0.35pp.

We finally examine the potential concern of reverse causality by instrumenting the contemporaneous IT-intensive share between 1990 and 2015 with the historical 1980 IT-intensive share. Our identifying variation emerges from the fact that counties with greater exposure to IT-intensive jobs in 1980 are more likely to experience an expansion during the 1990s technology boom. Our exclusion restriction, however, requires that unobserved shocks to contemporaneous manufacturing shares are uncorrelated with historical variation in the 1980 share. Since areas vary in unobserved ways, we insert demographic controls from 1980, including the share of college graduates, males, and whites, on top of our usual contemporaneous controls. Column 5, therefore, provides an estimate that is remarkably consistent with our baseline of a 0.37pp decline in the the manufacturing employment share.[17] At the very least, the robustness of our estimates shows that contemporaneous factors, like China's entry onto the global stage (Autor et al., 2013a), are unlikely to reverse the main result.

[INSERT TABLE 3 HERE]

[17]Failing to control for these 1980 demographic shares produces a downwards biased estimate of -0.67 since areas with higher employment shares of IT jobs in 1980 are likely positively selected and, therefore, also have lower manufacturing employment shares. While one might be tempted to control for the 1980 manufacturing share, this would produce a biased estimate since we would be instrumenting with a variable that endogenously affects our hypothetical control; Angrist and Pischke (2009) refer to it as "over controlling".

4 IT-task intensity and Labor Substitution

The comovement of price and quantities of different types of labor inputs conveys information about the extent to which IT-task intensive jobs can substitute for other jobs. In what follows we examine this comovement through the lens of a production structure featuring imperfect substitution between heterogeneous inputs. In this way we are able to estimate a well-defined elasticity of substitution between IT and non-IT labor. The analysis allows for this elasticity to differ across industries (manufacturing and services). We posit that output (value added) in industry i, denoted Φ_{it}, is generated using aggregate IT and non-IT efficiency-weighted labor services, denoted L_{it}^{IT} and L_{it}^{NIT}:

$$\Phi_{it} = \left[\alpha_{it}\left(L_{it}^{IT} z_{it}^{IT}\right)^{\nu_i} + (1-\alpha_{it})\left(L_{it}^{NIT} z_{it}^{NIT}\right)^{\nu_i}\right]^{\frac{1}{\nu_i}} \tag{2}$$

where z denotes task-specific productivity shocks. The random variables $z_{it}^{j} = \exp(\varepsilon_{it}^{j})$ are independent across industries i and task groups j, with $\varepsilon_{it}^{j} \sim iid(0, \sigma_{ij}^{2})$.[18] The price of each industry $i's$ output is p_{it} and can be thought of as the marginal value of that industry's output. Equation 2 captures the fact that there is imperfect subsitutability of IT and non-IT labor services. This subsitutability is allowed to vary across industries. The efficiency-weighted inputs are the product of total hours worked (raw hours) and efficiency per hour worked (quality):

$$L_{it}^{IT} = H_{it}^{IT} E_{it}^{IT}, \quad L_{it}^{NIT} = H_{it}^{NIT} E_{it}^{NIT} \tag{3}$$

where α_{it} denotes the factor share of IT-intensive labor services of workers in industry i. This factor share can be interpreted both as factor intensity and as the level of factor augmenting technical progress. The restriction that IT and non-IT input shares sum to 1 is made for convenience and is not strictly necessary. H_{it} denotes total work hours in industry i and period t, while E_{it} denotes average efficiency per hour worked in a given industry-time bin. Taking derivatives, the marginal products of *effective* labor services in industry i are the wages per efficiency unit (shadow

[18] We purposefully abstract from modeling capital for various reasons. First, our empirical analysis focuses on the ratio of marginal products of IT and non-IT labor, thereby allowing us to partial out capital. In this sense our empirical approach implicitly accommodates alternative specifications of capital across industries and education groups (including a simple Cobb-Douglas technology in capital and aggregate labor). Second, our primary object of interest is added value produced using IT and non-IT labor inputs. We can capture the degree of subsitutability of labor inputs in added value across industries and education groups without making strong assumptions about the role of capital. Finally, modeling capital is complex and beyond the scope of our analysis. In fact, we know of no obvious way of measuring and linking IT-specific capital and IT-intensive labor inputs.

prices) $\omega^{IT}$and ω^{NIT}. That is, the marginal products of IT and non-IT-intensive labor services can be written:

$$p_{it}MP_{it}^{IT} = \omega_{it}^{k} = \frac{\partial \Phi_{it}}{\partial L_{it}^{IT}} = \Phi_{it}^{1-\nu_i}\left[\alpha_{it}(L_{it}^{IT})^{\nu_i-1}(z_{it}^{IT})^{\nu_i}\right]$$

$$p_{it}MP_{it}^{NIT} = \omega_{it}^{NIT} = \frac{\partial \Phi_{it}}{\partial L_{it}^{NIT}} = \Phi_{it}^{1-\nu_i}\left[(1-\alpha_{it})\,(L_{it}^{NIT})^{\nu_i-1}(z_{it}^{NIT})^{\nu_i}\right]$$

The ratio of these marginal products implies the following optimality condition for each industry i:

$$\frac{\omega_{it}^{IT}}{\omega_{it}^{NIT}} = \left(\frac{\alpha_{it}}{1-\alpha_{it}}\right)\left(\frac{L_{it}^{IT}}{L_{it}^{NIT}}\right)^{\nu_i-1}\left(\frac{z_{it}^{IT}}{z_{it}^{NIT}}\right)^{\nu_i} \tag{4}$$

Working with ratios means that we can ignore the industry-specific output prices p_{it}. Taking logarithms $\omega_g^k = \ln(w_g^k)$ and time-differences delivers the following equation:

$$\Delta \ln(\frac{\omega_{it}^{IT}}{\omega_{it}^{NIT}}) = \Delta \ln\left(\frac{\alpha_{it}}{1-\alpha_{it}}\right) + (\nu_i - 1)\Delta \ln\left(\frac{L_i^{IT}}{L_i^{NIT}}\right) + \nu\Delta\left(\frac{\varepsilon_{it}^{IT}}{\varepsilon_{it}^{NIT}}\right) \tag{5}$$

The ratio $(\omega_{it}^{IT}/\omega_{it}^{NIT})$ denotes the relative return to IT-intensive tasks (IT premium) per efficiency unit, while (L_i^{IT}/L_i^{NIT}) is the IT premium in effective labor. Neither ratio is directly observable in data. Multiplying both sides of equations 4 and 5 by, respectively, $\ln(L_{it}^{IT}/L_{it}^{NIT})$ or $\Delta\ln(L_{it}^{IT}/L_{it}^{NIT})$ allows one to express these relationships as equations linking ratios of total earnings to ratios of inputs. This change of variable is helpful because earnings are observable, requiring that we only obtain measures of efficiency-weighted labor services. We discuss these measurement issues below.

4.1 Empirical Approach

There are several approaches to estimating the elasticity of substitution between IT and non-IT jobs in Equation5. We overview the different alternatives and discuss their advantages and disadvantages. We restrict our sample to four decades (1980, 1990, 2000, 2013) when we estimate Equation 5 in log-levels and to three decades (1990, 2000, 2013) when we estimate equation 5 in first-differences. Our unit of analysis consists of a state (52) × education (college / non-college) × industry (manufacturing / service)× year (1980, 1990, 2000, 2013). For every such unit of analysis we generate weighted measures of each variable. That is, we group together workers with common observable characteristics and form a synthetic panel. Hence, each observation is at the level of a (industry×location×education×period) group $g \in G$, where G is the total number of bins.

Before discussing our empirical strategies, we begin by outlining the various identification challenges that are present in estimating our industry-specific elasticity, ν_i. First, time-invariant differences across location and skill groups may be correlated with effective labor services. For example, if college educated workers tend to work longer and earn more for reasons that are correlated with, for example, preferences, then we might confound variation in skill prices with labor services. Second, since efficiency-weighted labor services are unobserved, we must construct a measure that directly captures the group-specific skill price, so that we can deflate labor services accordingly, or alternatively we need to model the sorting of heterogeneous workers into different jobs so to obtain a composition-free measure of labor services. Third, endogenous supply-side responses to demand or productivity shocks might be correlated with input utilization. For example, if productivity increases, firms may adjust their demand for labor services. Fourth, unobserved shocks to the change in the relative IT task premium might be correlated with changes in the labor services premium. For example, if workers respond to changes in the price of skill by sorting into different jobs, then the composition of workers in different groups would change.

Next, we sequentially describe our solution to each of the different identificat ion problems. Fortunately, the first challenge can be easily addressed by first-differencing the equilibrium condition linking prices with labor services, as in equation 5. This removes time-invariant heterogeneity from our synthetic panel data set.

To address the second challenge, related to measurement, we introduce a quality-adjustment of labor services based on estimates of the relative IT task price. Using the definition of effective (efficiency-weighted) labor aggregates, within each group $g \in G$, we have

$$L_g^{IT} = H_g^{IT} E_g^{IT}, \quad L_g^{NIT} = H_g^{NIT} E_g^{NIT}.$$

Furthermore, wage bills (aggregate earnings) within each group g are defined as the products $\omega_g^{IT} L_g^{IT}$ and $\omega_g^{NIT} L_g^{NIT}$. To approximate the relative price ratios, $\ln(\omega_g^{IT}/\omega_g^{NIT})$, and labor services ratios, $\ln(L_{it}^{IT}/L_{it}^{NIT})$, we apply a variant of the approach of Autor et al. (2006) and Autor et al. (2008) by proceeding in two steps. First, we regress logged annual earnings on an indicator for whether the individual is in an IT job, conditional on a range of observed characteristics. Under the identifying assumption that unobserved shocks to earnings are uncorrelated with selection into IT and non-IT jobs across industries, the coefficient captures the skill price of IT jobs. Second, we use the estimated coefficient (normalized by the individual's annual hours worked) to deflate the earnings ratio and obtain a quality-adjusted measure of labor services.[19] While this approach has its limitations, it is preferable to using raw hours worked as a proxy for labor services, which would require assuming that all individuals are identical in terms of their efficiency.

The third issue has to do with endogeneity of input choices made by optimizingfirms. We address this by instrumenting for labor inputsusing historical shocks (specifically the growth in the IT employment premiumbetween 1970 and 1980 and the share of IT workers in 1970). These historical instruments are predetermined with respect to contemporaneous movements in bothlabor services and skill price premia. Hence, they help show that our elasticity estimates are not driven by mechanical concerns about contemporaneous input utilization decisions. We also experiment using lagged values of labor inputs as instruments, obtaining similar results.

The final identification concern relates to the possibility of endogenous sorting into jobs based on unobserved characteristics. If individuals move across sectors because of unobserved comparative advantage, then our skill price proxy will be biased. In particular, as more and more marginal workers flow out of manufacturing into services, the relative efficiency composition of workers in manufacturing will improve significantly, given its relatively small size. Ignoring these composition effects may cause us to underestimate the elasticity of substitution between IT and non-IT workers in the manufacturing sector. In other words, changes in labor services might reflect differences in the quality of workers within the sector because individuals are responding to changes in relative task prices. We address this non-random selection by estimating a model featuring endogenous selection due to unobserved comparative advantage, based on the original contribution of Adao (2016). For brevity, we relegate the description of this model to Appendix A.4, where we outline its structure, the key identifying assumptions and estimating equations.

4.2 Estimates of the Elasticity by Industry

Table 4 presents estimates of the elasticities for different specifications of the estimating equations for the baseline model, which does not explicitly account for selection on unobservables, as well as for specifications based on selection-adjusted measures of quantities and prices. We discuss

[19]The identifying assumption here is not innocuous and we discuss its limitations below. We control for demographics (race, marital status, gender, a quadratic in years of schooling, age bin fixed effects [20-29, 30-39, 40-49, 50+], weeks worked per year (over the last year) fixed effects, and logged weekly hours worked. The advantage of this approach is that hours worked and weeks worked potentially absorb some of the variation in individual earnings dispersion that is correlated with preferences for work and/or other possible confounders. We have also experimented with regressing the hourly wage on an IT job indicator, using it to deflate the earnings ratio. These two approaches give similar results.

selection-adjusted estimates in the following section. Estimates are reported separately for each industry (M=manufacturing; S=services), for both log-levels levels and first-difference specifications (in which case "Time Diff.(Δ)" is set to "Yes").We report OLS estimates as well as IV (in which case "Instrument" is set to "Yes"). Time-differenced specifications are preferable in the presence of time-invariant unobserved heterogeneity.[20]

We begin by discussing results based on a simple quality-adjusted measure of labor inputs (obtained by deflating earnings using a IT task price premium that varies by group $\times$ period). Columns 1 and 2 report OLS estimates for log-level specifications, suggesting an elasticity of substitution of 1.72 for manufacturing and 1.24 for services. When we estimate the relationship in first-differences (columns 3 and 4) we obtain elasticities of, respectively, 1.28 and 1.15. Industry-specific elasticities appear even more similar in columns 5 and 6, where we adopt an instrumental variable approach using historical growth in the employment premium between 1970 and 1980 and the IT share of workers in 1970.

Turning to the estimates based on the selection model, which corrects for both quality and non-random sorting over jobs, we find elasticities of 1.63 and 1.31, respectively, for the manufacturing and services sectors. To account for the concern that changes in labor inputs may be endogenously related with changes in skill prices, we again instrument using our pre-determined historical measures, obtaining elasticities of 1.77 and 1.56. The industry-specific elasticities estimated using the selection-adjusted model are statistically different from one another: the p-value is close to zero for the OLS and is 0.028 for the IV. The wider wedge between the two sectoral elasticities in the selection-adjusted model is consistent with systematic changes in average efficiency due to the reallocation of marginal workers from manufacturing into services. Finally, Table 8 in Appendix A.5 presents a set of elasticity estimates based on an alternative measure in which we define high IT intensity occupations those with intensity scores above the median.

[INSERT TABLE 4 HERE]

[20] We use a simultaneous equation estimator that allows us to (i) flexibly combine information for different decades, (ii) experiment with and without restricting parameters to be the same across decades, (iii) control for endogeneity through first-stage IV regressions, when necessary.

5 The Emergence of IT-Intensive Services and Solow's Paradox

Traditional arguments about the increasing pervasiveness of IT point towards greater productivity and automation in the work-place. However, as Acemoglu et al. (2014) clearly show for the U.S. manufacturing sector, there is little systematic evidence that IT-using industries have actually become more productive over time. While there is some evidence of faster growth in output per worker in IT-using industries, these differences are driven by a more rapid decline in labor employed in those industries and a relatively constant, or even declining, level of output (rather than higher output, as one might have expected). For these reasons, Acemoglu et al. (2014) conclude that Solow's famous paradoxical statement that "...you can see the computer age everywhere but in the productivity statistics" is far from being resolved. We revisit these issues through the lens of our model and make two contributions: (i) we use our measure of the IT employment share based on occupational heterogeneity to replicate Acemoglu et al. (2014) and compare our results with those obtained using their industrial heterogeneity measure, and (ii) we apply our IT measure to distinguish between labor productivity growth in high and low IT industries within both the manufacturing and services sectors.[21]

We begin by constructing the employment share of IT workers across each three-digit industry, comparing it with the baseline IT capital intensity measure in Acemoglu et al. (2014). The (weighted) correlation between these alternative measures of IT prevalence is 0.46 (see Figure 9).[22] One reason our measure differs from industry-based proxies is that we focus on the underlying tasks that workers perform at a five-digit occupation level, rather than measuring capital assets as in prior work (e.g., Stiroh (2002)). While we do not argue that looking at a measure of IT-using labor is better than a measure of IT-using capital, we simply point out that they capture different features of reality. Using our employment share of IT by industry, we subsequently estimate regressions identical to Acemoglu et al. (2014) at the three-digit industry $\times$ year level:

[21] To put the benefit of our task-based measure into context, Acemoglu et al. (2014) remark that: "A second category of explanation for these unexpected results is that our measure of IT investment, constructed by averaging computer investment data from 1977–2007, misses the mark."

[22] We also implemented a different comparison by partitioning industries into high and low IT based on our IT-intensity measure, as well as the measures from Acemoglu et al. (2014) based on computer investment, the SMT dummy, and computer investment between 1977-82. The results imply that the first measure matches our high and low IT industries in 72% of the cases, the second measure only matches ours in 48% of the cases, and the third measure matches ours in 64% of the cases.

$$\log Y_{jt} = \gamma_j + \delta_t + \sum_t \beta_t \times IT_j + \epsilon_{jt} \tag{6}$$

where Y denotes our outcome variable (e.g., the real value of shipments or real output), IT denotes a measure of the IT intensity, γ denotes three-digit industry fixed effects, and δ denotes year fixed effects. The primary coefficients of interest in Equation 6 are the β^t values. We normalize the coefficient on the IT variable to zero in the base year β^0. Hence the series $\{\beta^1, \beta^2, ...\}$ may be read as the level of the coefficient on IT in each subsequent year, relative to the base year. These coefficients characterize the deviations from trend in the return to IT intensity across industries: hence, only their time variation (rather than their absolute value) is of interest for our analysis.

As mentioned above, we use the baseline measure of IT intensity taken from Acemoglu et al. (2014), which is the share of investment in computers (relative to total investment), together with our employment share of IT, which denotes the share of workers in IT-intensive jobs within a given industry.

[INSERT FIGURE 9 HERE]

Figure 10 plots the estimated δ_t coefficients for each year. We follow Acemoglu et al. (2014) as closely as possible by crosswalking industries from our CPS micro-data to the three-digit industry level, which we match into their SIC codes. We also drop the sample of industries that falls within their "medium computer intensity" measure since those industries expand mechanically over time due to the increasing pervasiveness of computers in the economy.[23] The blue line in Figure 10 replicates their main result that IT use at the industry level is negatively associated with real shipments over time. However, the red line in Figure 10 illustrates a near opposite pattern when using our measure of IT intensity (correlation = -0.79)—that is, rising IT intensity is associated to positive changes in the real value of shipments. As we discussed earlier, one advantage of our measure is that it captures variation in the utilization of different tasks within each industry, which is important since IT content is inherently occupation-specific.

Why do these measures produce such heterogeneous productivity profiles? Our measure focuses on tasks (and associated human capital among the labor force) in IT jobs, whereas Acemoglu et al. (2014) focus on an industry-level measure of physical IT capital. Hence these two indices are capturing systematically different phenomena. Given our earlier evidence that the manufacturing

[23] These industries in the "medium intensity" range of the classification include: SIC 3570-3579, SIC 3660-3669, and SIC 3670-3679. They also consider a e narrower and a broader definition from Houseman et al. (2013).

sector exhibits greater subsitutability between IT and non-IT workers, our analysis suggests that the impact of a marginal unit of IT intensive labor may be strong exactly in those industries that invested more in physical capital. In sum, these differences underscore that the measurement strategy matters.

[INSERT FIGURE 10 HERE]

Motivated by these differences between our two measures, we next examine how the IT employment share affects our understanding of productivity growth in the manufacturing versus services sectors. That is, we now estimate Equation 6 using logged real output (deflated with 2009 price indices) as the outcome variable and the IT employment share as the IT intensity index, plotting the estimated δ^t coefficients separately for the manufacturing and services sectors.[24] Figure 11 highlights two important observations. First, and perhaps most importantly, whereas IT intensity explains a relatively constant share of real output productivity over the past 40 years in the manufacturing sector, IT intensity is growing rapidly in explaining the surge of productivity in the services sector. Second, the positive effect of IT intensity in the services sector has continued unabated after the year 2000. These results are consistent with our early evidence that the bulk of productivity growth took place within high IT services (Figure 5).

[INSERT FIGURE 11 HERE]

6 Conclusion

Recent debates about the role of technology and automation for labor productivity across jobs have highlighted the need for new and alternative ways to describe and quantify the emergence of information technologies in production. Important aspects of the growth of information technology (IT) include its impact on the labor market (Autor et al., 1998), the organization of firms (Bresnahan et al., 2002), productivity (Brynjolfsson and Hitt, 2003), and managerial practices (Bloom et al., 2012). However, much less is known about how the emergence of IT intensive jobs has influenced structural transformation—that is, productivity, wages and employment changes in manufacturing and services.

We develop a new index of occupation-level IT intensity that draws from O*NET measures of task, skill and work environment characteristics. After introducing our index and classifying

[24] We use real output since it most closely matches the real value of shipments used in Acemoglu et al. (2014).

five-digit occupations as either high or low IT intensity, we document three stylized facts. First, the bulk of the increases in productivity since 1950 has been concentrated in high IT industries. In particular, the high IT services sector has experienced a fivefold growth in value added per hour worked between 1950 and 2010. Second, we document how technological change has shaped the distribution of occupation-level returns and employment. We show that IT-intensive jobs exhibit large and growing earnings and employment premia since 1970. Despite a flattening of the earnings and employment premia for college and cognitive/non-routine jobs after 2000, premia have continued rising in IT intensive jobs. Third, we show that increases in the IT employment share are associated with declines in the manufacturing employment share, consistent with the view that the surge in IT productivity has accelerated structural transformation.

We subsequently use our task-based measure to estimate the elasticity of substitution between IT and non-IT jobs in production. Under our preferred specification, we find that the elasticity of substitution between IT and non-IT labor inputs is 1.6 in the manufacturing sector and 1.3 in the services sector.

We conclude by using our IT intensity measure to examine whether the emergence of IT-intensive jobs is associated with productivity growth. Building on Acemoglu et al. (2014), we find that increases in the prevalence of IT intensive jobs are associated with growth in the real value of shipments and output at the industry level.

Our results are not meant to provide a definitive conclusion on the relationship between occupation-level IT intensity and structural transformation. Rather, we present initial evidence on this relationship to discipline and motivate additional research questions about the ways in which technological changes affect productivity and inequality patterns across sectors. In particular, we have shown that the emergence of IT intensive jobs has had a major impact on the employment structure and on the distribution of wages both within and across sectors. These changes may in principle account for a non-trivial part of the increasing polarization across jobs documented by Autor and Dorn (2013). We leave further analysis of these implications for future work.

References

Acemoglu, D., D. Autor, D. Dorn, G. H. Hanson, and B. Price (2014): "Return of the solow paradox? IT, productivity, and employment in US manufacturing," *American Economic Review: Papers & Proceedings*, 104, 394–399.

Acemoglu, D. and D. H. Autor (2011): "Skills, tasks and technologies: Implications for employment and earnings," *Handbook of Labor Economics*, 4b, 1043–1171.

Acemoglu, D., D. H. Autor, and D. Lyle (2004): "Women, war and wages: The effect of female labor supply on the wage structure at midcentury," *Journal of Political Economy*, 112, 497–551.

Acemoglu, D. and V. Guerrieri (2008): "Capital deepending and nonbalanced economic growth," *Journal of Political Economy*, 116, 467–498.

Acemoglu, D. and P. Restrepo (2017a): "Robots and jobs: Evidence from US labor markets," *Working paper*.

——— (2017b): "The race between man and machine: Implications of technology for growth, factor shares and employment," *Working paper*.

Adao, R. (2016): "Worker heterogeneity, wage inequality, and international trade: Theory and evidence from Brazil," *Working paper*.

Angrist, J. D. and J.-S. Pischke (2009): "Mostly harmless econometrics: An empiricist's companion," *Princeton University Press*.

Atalay, E., P. Phongthiengtham, S. Sotelo, and D. Tannenbaum (2017): "The evolving U.S. occupational structure: A textual analysis," *Working paper*.

Autor, D. (2017): "How long has this been going on?" *NBER Conference on Research and Income in Wealth*.

Autor, D. and D. Dorn (2013): "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market," *American Economic Review*, 103, 1553–1597.

Autor, D. H. (2013): "The "task approach" to labor markets: An overview," *NBER working paper*.

Autor, D. H., D. Dorn, and G. H. Hanson (2013a): "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 103, 2121–2168.

——— (2015): "Untangling trade and technology: Evidence from local labour markets," *Economic Journal*, 125, 621–646.

Autor, D. H., Katz, and M. J. Handel (2013b): "Putting tasks to the test: Human capital, job tasks and wages," *Journal of Labor Economics*, 31, S59–S96.

Autor, D. H., L. F. Katz, and M. S. Kearney (2006): "The Polarization of the U.S. Labor Market," *American Economic Review*, 96, 189–194.

——— (2008): "Trends in U.S. wage inequality: Revising the revisionists," *Review of Economics and Statistics*, 90, 300–323.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998): "Computing inequality: Have computers changed the labor market?" *Quarterly Journal of Economics*, 113, 1169–1213.

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The skill content of recent technological change: An empirical exploration," *Quarterly Journal of Economics*, 118, 1279–1333.

Barany, Z. and C. Siegel (2017): "Job polarization and structural change," *American Economic Journal: Macroeconomics*, forthcoming.

Baumol, W. J. (1967): "Macroeconomics of unbalanced growth: Anatomy of an urban crisis," *American Economic Review*, 57, 415–426.

Beaudry, P., D. A. Green, and B. M. Sand (2014): "The declining fortunes of the young since 2000," *American Economic Review*, 104, 381–386.

——— (2016): "The Great Reversal in the demand for skill and cognitive tasks," *Journal of Labor Economics*, 34, S199–S247.

Beckhusen, J. (2016): "Occupations in information technology," *American Community Survey Reports*.

Bloom, N., R. Sadun, and J. Van Reenen (2012): "Americans do IT better: US multinationals and the productivity miracle," *American Economic Review*, 102, 167–201.

Boppart, T. (2014): "Structural change and the Kaldor Facts in a growth model with relative price effects and non-Gorman preferences," *Econometrica*, 82, 2167–2196.

Bresnahan, T., E. Brynjolfsson, and L. Hitt (2002): "Information technology, workplace organization and the demand for skilled labor: firm-level evidence," *Quarterly Journal of Economics*, 117, 339–376.

Bryne, D. M., J. G. Fernald, and M. B. Reinsdorf (2016): "Does the United States have a productivity slowdown or a measurement problem?" *Brookings Papers on Economic Activity*, 109–182.

Bryne, D. M., S. Oliner, and D. E. Sichel (2013): "Is the information technology revolution over?" *International Productivity Monitor*, 25, 20–36.

Brynjolfsson, E. and L. M. Hitt (2003): "Computing productivity: Firm-level evidence," *Review of Economics and Statistics*, 85, 793–808.

Brynjolfsson, E. and A. McAfee (2011): "Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy," *Lexington, MA: Digital Frontier Press*.

Brynjolfsson, E., D. Rock, and C. Syverson (2017): "Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics," *NBER working paper*.

Buera, F. J. and J. P. Kaboski (2009): "Can traditional theories of structural transformation fit the data?" *Journal of the European Economic Association*, 7, 469–477.

——— (2012): "The rise of the service economy," *American Economic Review*, 102, 2540–2569.

Buera, F. J., J. P. Kaboski, and R. Rogerson (2015): "Skill biased structural change," *NBER Working Paper*.

Card, D. and T. Lemieux (2001): "Can falling supply explain the rising return to college for younger men? A cohort based analysis," *Quarterly Journal of Economics*, 116, 705–746.

Charles, K. K., E. Hurst, and M. J. Notowidigdo (2016): "The masking of the decline in manufacturing employment by the housing buble," *Journal of Economic Perspectives*, 30, 179–200.

Comin, D. A., D. Lashkari, and M. Mestieri (2015): "Structural change with long-run income and price effects," *Econometrica, R&R*.

Deming, D. J. (2017): "The growing importance of social skills in the labor market," *Quarterly Journal of Economics*, 132, 1593–1640.

Duernecker, G., B. Herrendorf, and A. Valentinyi (2017): "Structural change within the service sector and the future of Baumol's disease," *Working paper*.

Eden, M. and P. Gaggl (2017): "On the welfare implications of automation," *Working paper*.

Feldstein, M. (2015): "The United States underestimates growth," *Wall Street Journal*.

Frey, C. and M. A. Osborne (2013): "The future of employment: How susceptible are jobs to computerisation?" *Working paper*.

Gollin, D., S. Parente, and R. Rogerson (2002): "The role of agriculture in development," *American Economic Review*, 92, 160–164.

Griliches, Z. (1969): "Capital-skill complementarity," *Review of Economics and Statistics*, 51, 465–468.

Hagerty, M. R. and K. C. Land (2007): "Constructing summary indices of quality of life: A model for the effect of heterogeneous importance weights," *Sociological Methods Research*, 35, 455–496.

Harrigan, J., A. Reshef, and F. Toubal (2016): "March of the techies: Technology, trade, and job polarization in France, 1994-2007," *Working paper*.

Hatzius, J. and K. Dawsey (2015): "Doing the sums on productivity paradox v2.0," *Goldman Sachs U.S. Economics Analyst*, 15/30.

Hecker, D. (2005): "High-technology employment: A NAICS-based update," *Monthly Labor Review*.

Herrendorf, B., C. Herrington, and A. Valentinyi (2015): "Sectoral technology and structural transformation," *American Economic Journal: Macroeconomics*, 7, 104–133.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2013): "Two perspectives on preferences and structural transformation," *American Economic Review*, 103, 2752–2789.

——— (2014): "Growth and structural transformation," *Handbook of Economic Growth*, 2B, 855–941.

Houseman, S. N., T. Bartik, and T. Sturgeon (2013): "Measuring manufacturing: Problems of interpretation and biases in the U.S. statistics," *Upjohn Institute*.

Juhn, C., K. M. Murphy, and B. Pierce (1993): "Wage inequality and the rise in returns to skill," *Journal of Political Economy*, 101, 410–442.

Kambourov, G. and I. Manovskii (2013): "A cautionary note on using (March) CPS and PSID data to study worker mobility," *Macroeconomic Dynamics*, 17, 172–194.

Katz, L. F. and K. M. Murphy (1992): "Changes in relative wages, 1963-1987: Supply and demand factors," *Quarterly Journal of Economics*, 107, 35–78.

Kongsamut, P., S. Rebelo, and D. Xie (2001): "Beyond balanced growth," *Review of Economic Studies*, 68, 869–882.

Kuznets, S. (1973): "Modern economic growth: Findings and reflections," *American Economic Review*, 63, 247–258.

Lemieux, T. (2006): "Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?" *American Economic Review*, 96, 461–498.

Mokyr, J. (2014): "Secular stagnation? Not in your life," *London: CEPR Press.*

Ngai, L. R. and C. A. Pissarides (2007): "Structural change in a multisector model of growth," *American Economic Review*, 97, 429–443.

Roy, A. (1951): "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 3, 135–146.

Schott, P. (2004): "Across-product versus within-product specialization in international trade," *Quarterly Journal of Economics*, 119, 647–678.

Stiroh, K. J. (2002): "Information technology and the U.S. productivity revival: What do the industry data say?" *American Economic Review*, 92, 1559–1576.

Syverson, C. (2017): "Challenges to mismeasurement explanations for the US productivity slowdown," *Journal of Economic Perspectives*, 31, 165–186.

Tambe, P. and L. M. Hitt (2012): "The productivity of information technology investments: New evidence from IT labor data," *Information Systems Research*, 23, 599–617.

Valletta, R. G. (2016): "Recent flattening in the higher education wage: premium: Polarization, skill downgrading, or both?" *Education, skills, and technical change: Implications for future U.S. GDP growth.*

7 Figures and Tables

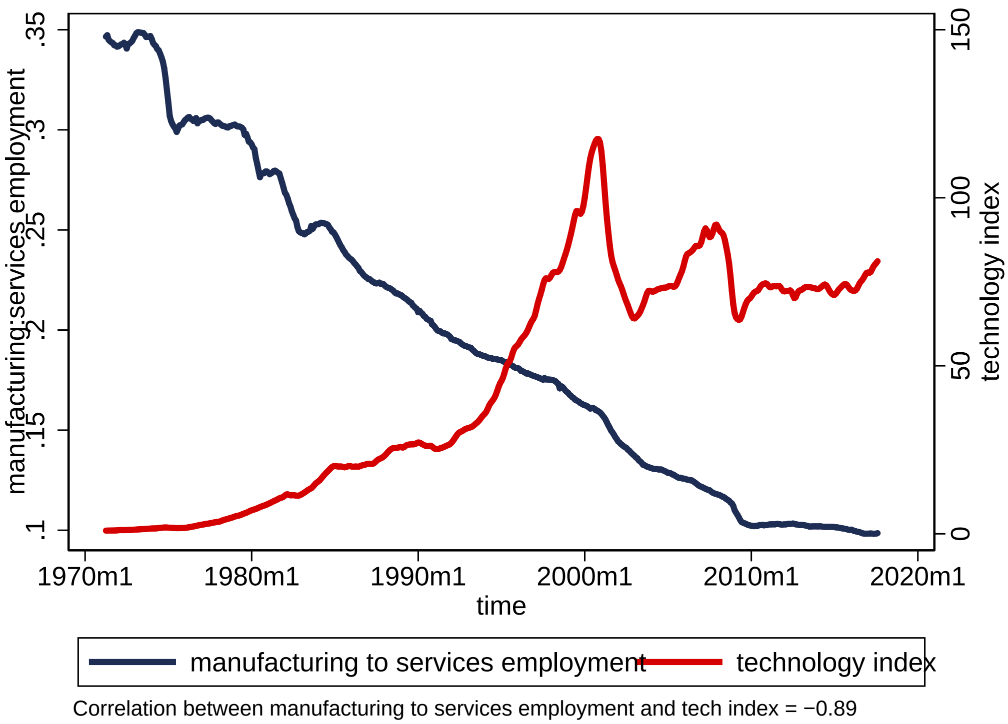


Figure 1: Structural Transformation and the Rise of Information Technology

Notes.–Sources: St. Louis Federal Reserve Board. The figure plots the ratio of manufacturing employment to services employment in the US economy vs a well-known technology index. The Federal Reserve Bank of San Francisco collects the technology index ("the tech pulse index") and is defined (from the FRED site) as follows: "a coincidence index of activity in the U.S. information technology sector. The index interpreted as the health of the tech sector". The indicators they use to construct the index are: investment in IT goods, consumption of personal computers and software, employment in the IT sector, industrial production of the technology sector, and shipments by the technology sector.

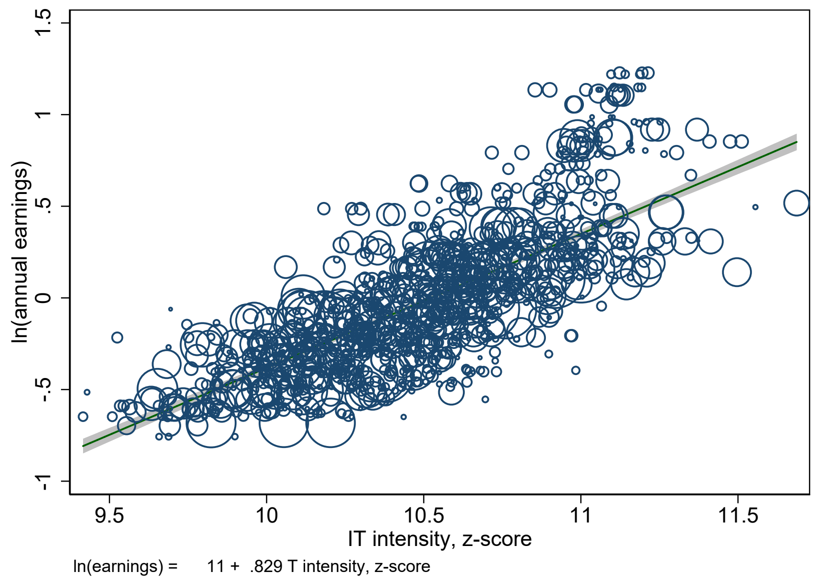


Figure 2: Employment Share and Earnings in Information Technology Jobs

Notes.–Sources: O*NET and Census Bureau, 1980-2015. The figure plots logged annual earnings (weighted by the Census sample weights) against the occupationj-level standardized IT intensity (weighted by employment from the BLS) at the five-digit aggregation level. Earnings are deflated using the 2010 personal consumption expenditure index. The sample is restricted to workers earning over $5,000/year, $2/hour, and working 500 hours/year. The figure shows a significant positive relationship between IT intensity and earnings in the cross-section of occupations.

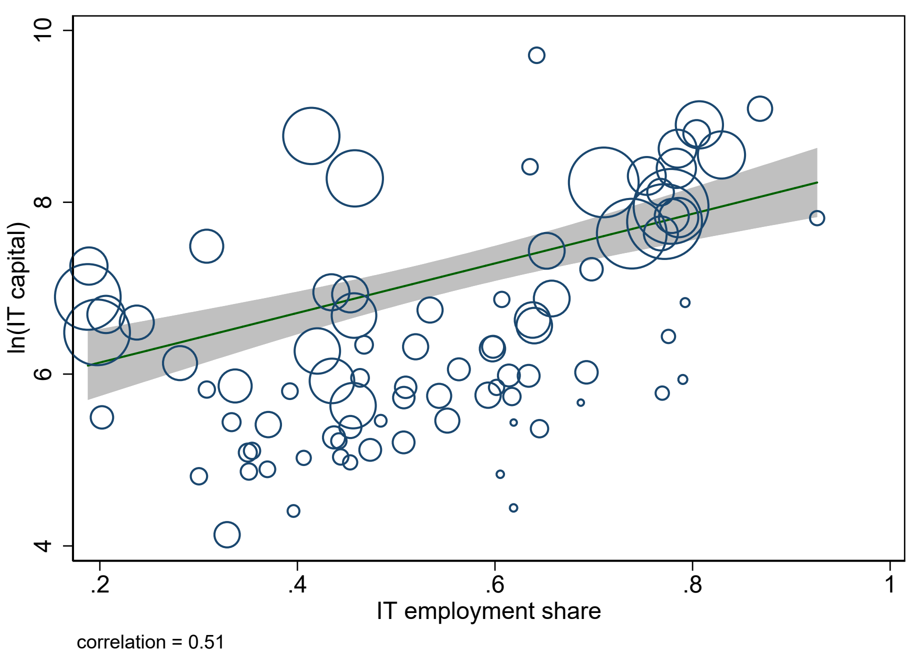


Figure 3: Validating Occupation Information Technology (IT) with Capital IT

Notes.–Sources: Census Bureau, O*NET, Bureau of Economic Analysis, 2000-2015. The figure plots the IT employment share at a three-digit industry level versus logged capital IT expenditures (weighted by employment from the BLS). The figure documents a positive 0.51 correlation between capital and labor IT inputs in the cross-section of three-digit industries (weighted by employment).

Table 1: Descriptive Statistics on High and Low Information Technology Workers

	pooled		1970s		1980s		1990s		2000s		2010-2015	
	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low
Demographics												
family size	2.9	3.1	3.4	3.4	3.0	3.1	2.9	3.1	2.9	3.0	2.9	3.0
number of children	1.0	0.9	1.4	1.3	1.0	1.0	0.9	0.9	0.9	0.9	0.9	0.8
age	41.5	39.8	40.2	40.3	38.7	37.6	39.1	37.7	40.7	39.2	42.7	40.7
male	0.58	0.53	0.76	0.59	0.67	0.54	0.59	0.53	0.57	0.52	0.54	0.52
married	0.65	0.54	0.80	0.72	0.72	0.64	0.68	0.59	0.66	0.55	0.61	0.46
white	0.83	0.76	0.95	0.87	0.92	0.86	0.88	0.80	0.84	0.76	0.79	0.72
black	0.08	0.13	0.04	0.12	0.06	0.12	0.07	0.12	0.07	0.12	0.09	0.14
schooling	14.5	12.2	13.4	10.6	14.1	11.7	14.4	12.3	14.5	12.4	14.9	12.6
college	0.52	0.14	0.33	0.04	0.41	0.08	0.47	0.11	0.51	0.13	0.58	0.18
Work												
hours worked	2105	1933	.	.	2017	1878	2058	1923	2110	1948	2123	1933
earnings	54658	30395	44308	27390	43121	27121	47347	29010	54023	31859	58486	30820
hourly wage	27.1	15.9	.	.	21.4	14.7	23.7	15.1	26.0	16.4	28.3	15.8
manufacturing	0.23	0.27	0.33	0.40	0.30	0.37	0.27	0.29	0.24	0.26	0.20	0.21
services	0.69	0.69	0.57	0.51	0.61	0.59	0.64	0.67	0.69	0.70	0.73	0.75
Observations	8123943	11141496	377515	719532	1408330	2183323	1617786	2372254	1723147	2297422	1546484	1806519

Notes.–Sources: Census Bureau and O*NET, 1970-2015. The table reports summary statistics associated with major demographic and labor market variables for high and low information technology (IT) workers by decade in our micro-data. The variables "manufacturing" and "services" refer to the employment share of workers. We refer to high IT workers as those classified as above the median based on our five-digit IT intensity z-score. Earnings are deflated using the 2010 personal consumption expenditure index and the sample is restricted to workers earning over \$5,000/year, \$2/hour, and working 500 hours/year. The manufacturing sector includes both the standard NAICS 31-33 codes on durables and non-durables and both agriculture and mining. The services sector includes finance, insurance, real estate, business services, wholesale and retail trade, utilities, transportation, personal and other services.

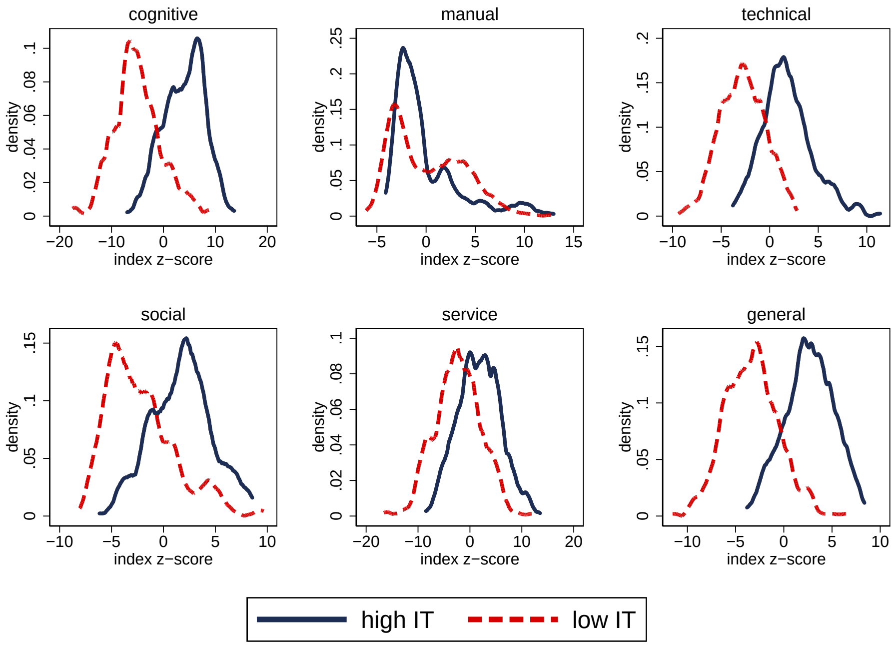


Figure 4: Distribution of Tasks Across Technology & Information Services Jobs

Notes.–Sources: Occupation and Employment Statistics (national) and O*NET, 2004-2014. The figure plots the distribution of the six major categories of tasks between high and low information technology intensive jobs. The skill groups are as follows: (1) cognitive tasks (decision making, learning strategies, listening, learning, problem solving, coordination, and critical thinking), (2) manual (repairs, equipment maintenance, equipment selection, installation, instruction), (3) technical (programming, quality control analysis, systems analysis, systems evaluation, technology design), (4) social (persuasion, social, speaking, negotiation), (5) service (management of financial resources, of material resources, of personnel resources, monitoring, service, operations control, operations monitoring, operations analysis, troubleshooting), and (5) general (math, writing, time management, reading, science). The ONET skill data is available from 2010-2014 and is made to have a mean zero and variance of 1. Occupations are harmonized to 2010 SOC codes.

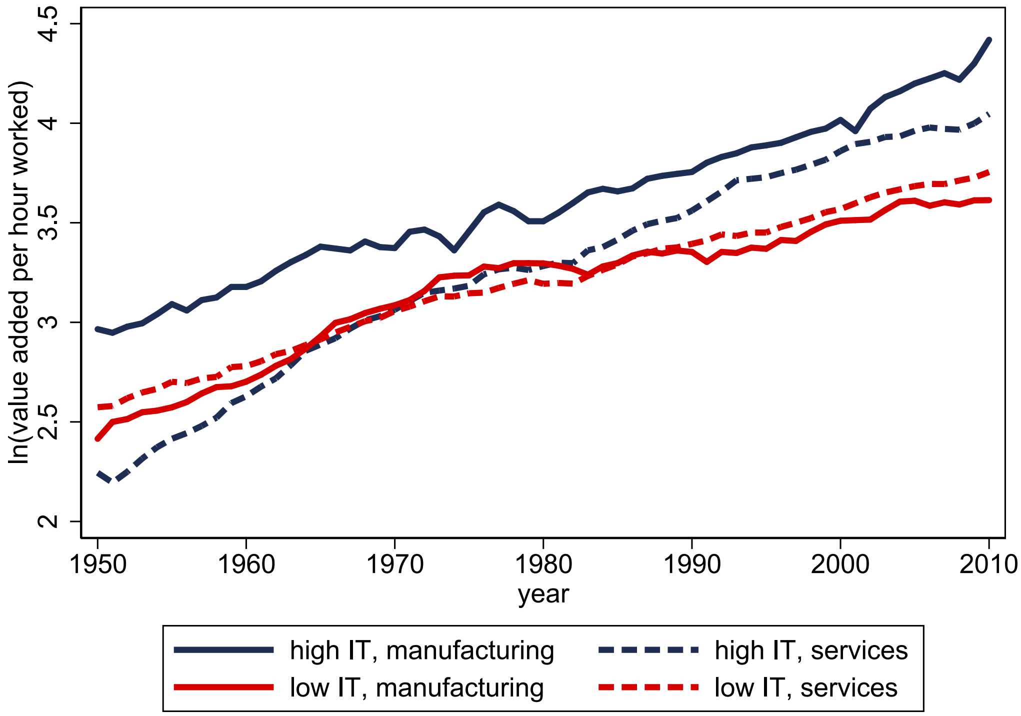


Figure 5: Value Added per Hour Worked in Manufacturing and Services, 1950-2010

Notes.–Sources: USA KLEMS, 1950-2010. The figure plots sectoral average annual value added per hour worked in constant thousands of dollars (deflated using the 2009 GDP deflator from the St. Louis Fed) for high and low IT intensity manufacturing and services sectors. Values within each high/low IT and manufacturing/services category are averaged based on the long-run employment counts between 1950 and 2010. Industries are classified as IT intensive if the weighted average z-score of information technology intensity within three-digit NAICS codes is greater than zero. We construct these weighted average scores using a sample of Census micro-data for 2000, 2006-2008, and 2013-2015 matched with IT intensities at a five-digit level as implemented in the baseline definition. The figure shows that logged value added per hour worked has grown primarily in the high IT manufacturing and services sector with the bulk of the growth driven by high IT services.

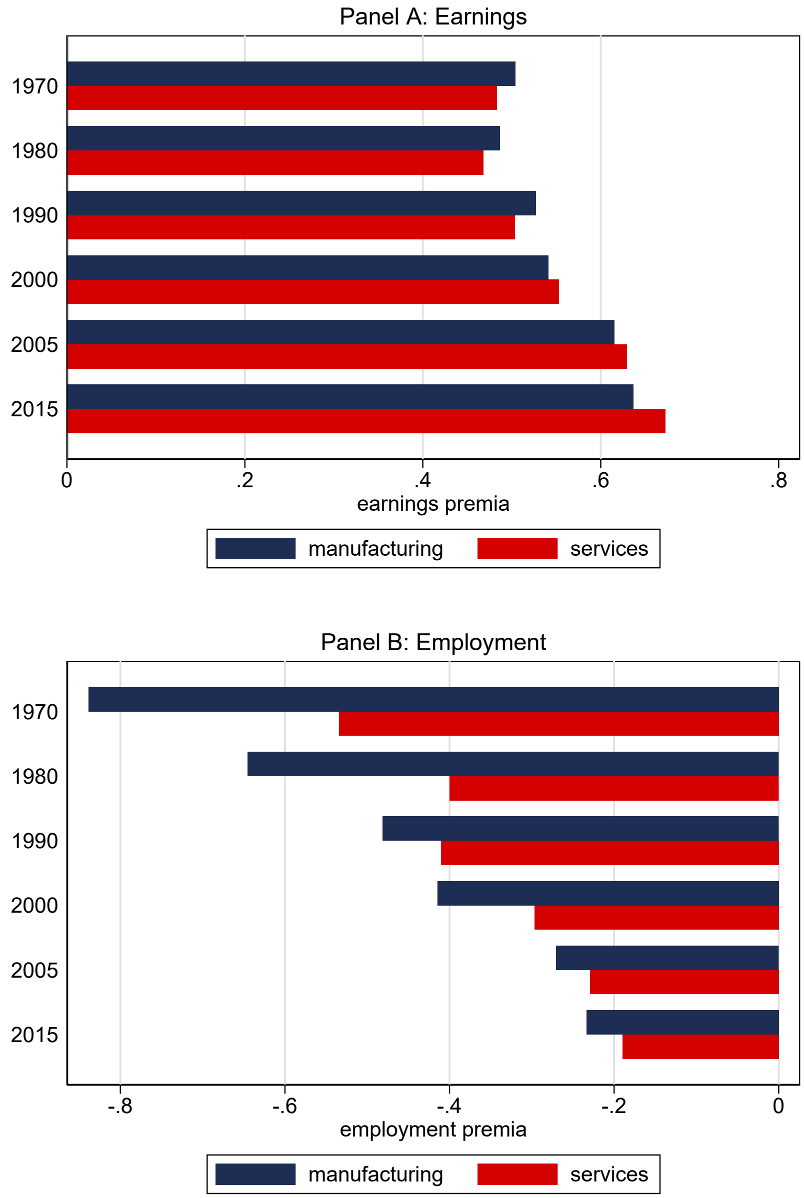


Figure 6: Earnings and Employment Premia to Information Technology, 1970-2015

Notes.–Sources: Census Bureau and O*NET, 1970-2015. Panel A plots the logged (average) annual earnings among IT workers net of logged annual earnings among non-IT workers (weighted by Census sample weights to produce the average) in the manufacturing and services sector. Earnings is deflated using the 2010 personal consumption expenditures index and the sample is restricted to workers earning over $5,000/year, $2/hour, and working 500 hours/year. Panel B plots the logged total number of IT workers net of the logged total number of non-IT workers. IT workers are classified as those in an occupation with an IT intensity score above the median in the distribution of five-digit occupations. The figure shows the growing earnings premium and declining employment premium (i.e., growing employment share) of IT workers.

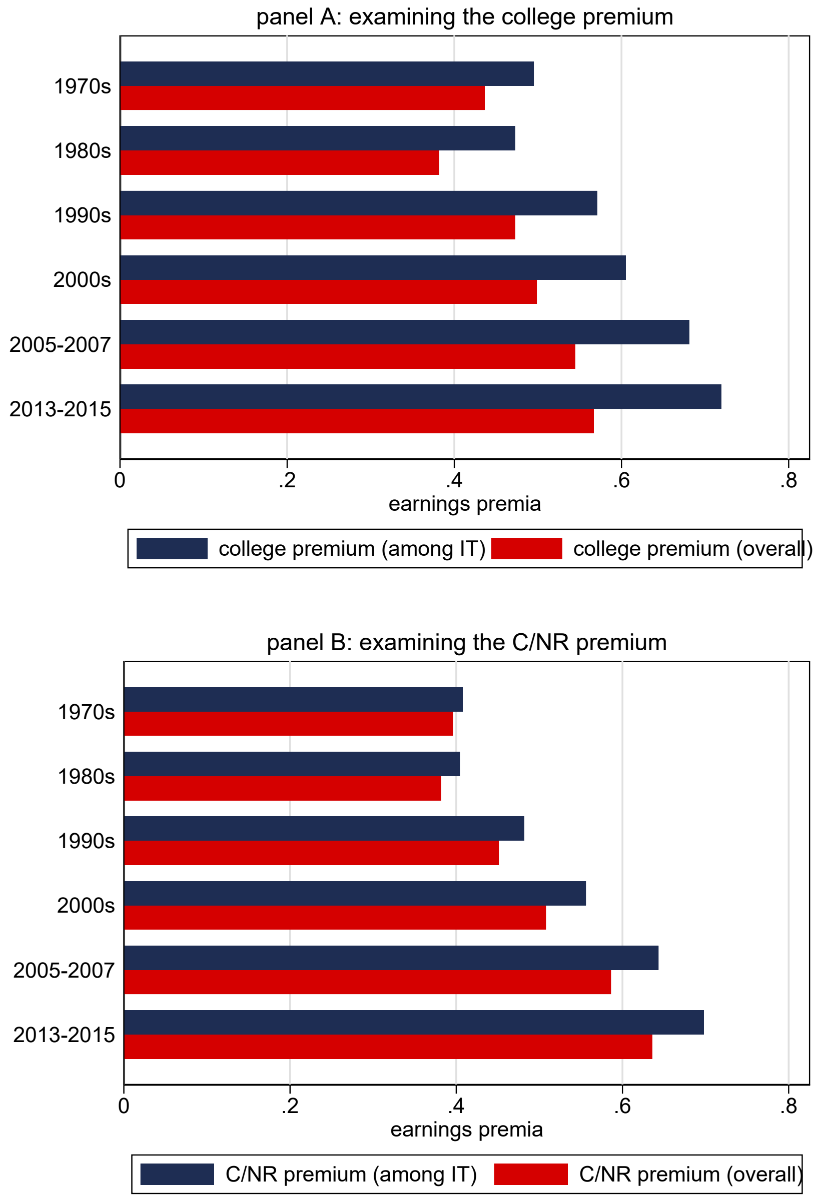


Figure 7: Examining Heterogeneity in the Demand for Skills in High and Low IT Jobs

Notes.–Sources: Census Bureau and O*NET, 1970-2015. Panel A plots, in blue, logged average earnings among IT intensive college workers net of logged earnings among non-college workers, together with, in red, logged earnings among college degree workers net of logged earnings among non-college degree workers. Panel B plots, in blue, logged earnings among high cognitive and non-routine workers net of logged earnings among low C/NR workers, together with, in red, logged earnings among high C/NR jobs. Earnings are deflated using the 2010 personal consumption expenditure index and the sample is restricted to workers earning over $5,000/year, $2/hour, and working 500 hours/year. All aggregations from the Census are produced using sample weights.

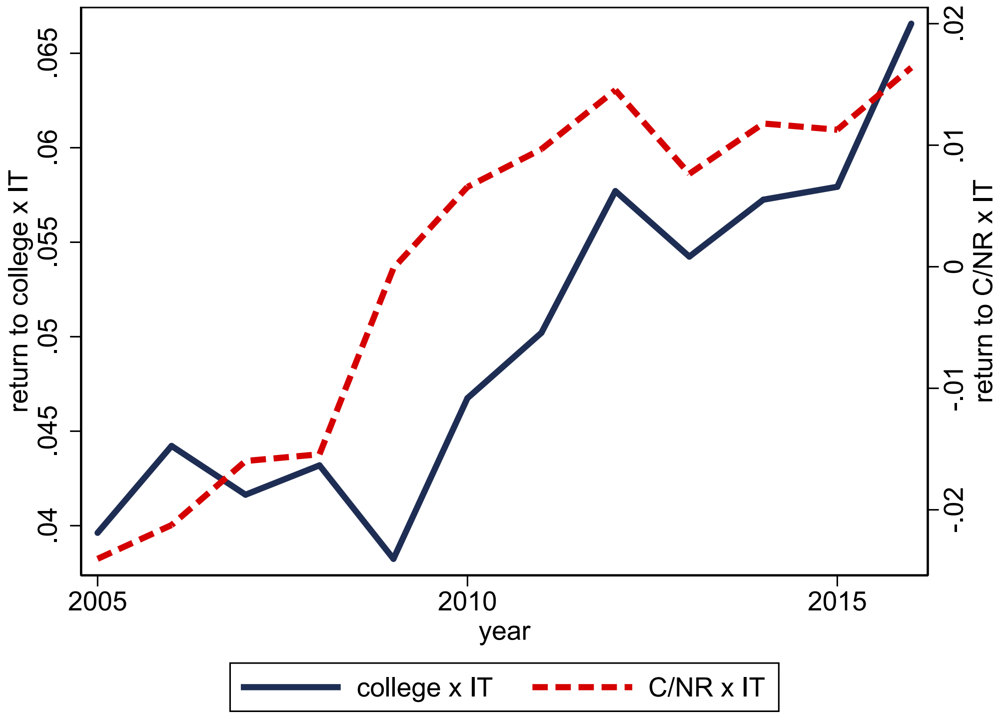


Figure 8: Return to IT and College / C/NR Skills, 2005-2016

Notes.–Sources: Census Bureau and O*NET, 2005-2016. Panel A plots the coefficient associated with the college × high IT status interactions from regressions of logged hourly wages (annual earnings divided by the product of weeks worked last year and average hours worked/week) on college attainment, high IT status, and their interaction separately by year, conditional on controls. Controls include: a quadratic in age, family size, race (white and black), marital status, and gender. Panel B implements an analogous comparison, but uses cognitive and non-routine skills classified at the five-digit occupation level based on whether the intensity is above or below the median. Earnings are deflated using the 2010 personal consumption expenditure index and the sample is restricted to workers earning over $5,000/year, $2/hour, and working 500 hours/year. Observations are weighted using sample weights and standard errors are clustered at the five-digit occupation level.

Table 2: Comparison of Baseline and Hecker (2005) Technology Measures

Dep. var. =	ln(annual earnings)				ln(annual hours)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IT (baseline)	.63***		.43***		.10***		.08***	
	[.06]		[.03]		[.02]		[.01]	
high-tech (hecker, 2005)		.55***		.34***		.09***		.05***
		[.04]		[.04]		[.01]		[.01]
R-squared	.18	.04	.34	.28	.02	.00	.07	.06
Sample Size	10379410	10379410	10379410	10379410	10457519	10457519	10457519	10457519
Controls	No	No	Yes	Yes	No	No	Yes	Yes

Notes.–Sources: Census Bureau, O*NET, 1970-2015. The table reports the coefficients associated with regressions of logged annual earnings and hours worked on an indicator for whether the person is in an IT intensive job using our baseline O*NET measure and an alternative from Hecker (2005), conditional on controls, which include the number of children, family size, age, gender, race, and fixed effects on education. Standard errors are heteroskedasticity-robust and observations are weighted by sample weights.

Table 3: The Rise of the Information Technology Employment Share and Structural Change

Dep. var. =	manufacturing employment share				
	(1)	(2)	(3)	(4)	(5)
IT share	-.67***	-.23**	-.38***	-.35***	-.37**
	[.06]	[.09]	[.06]	[.09]	[.17]
ln(wage)		.12***		.11***	.12***
		[.03]		[.03]	[.03]
male		.95***		.72***	1.07***
		[.07]		[.09]	[.14]
white		.14***		-.00	.23***
		[.02]		[.03]	[.03]
college		-.37***		-.23***	-.21**
		[.05]		[.06]	[.09]
R-squared	.22	.61	.92	.94	.57
Sample Size	1434	1434	1422	1422	1090
Controls	No	Yes	No	Yes	Yes
County FE	No	No	Yes	Yes	No
Year FE	No	No	Yes	Yes	No
Instrument?	No	No	No	No	Yes

Notes.–Sources: Census Bureau, 1980-2015. The table reports estimated coefficients obtained by regressing the manufacturing employment share at the county-level on the IT employment share, conditional on the logged hourly wage, share of males, share of whites, and share of college degree workers. There are roughly 250 counties in each year of our sample. Fixed effects are included for county and year. The instrumental variable specification in column (5) uses the pre-determined 1980 share of IT workers to predict future IT growth. The sample is restricted to workers earning over $5,000/year, $2/hour, and working at least 500 hours/year. Standard errors are clustered at the county-level.

	Quality-Adjusted Measures						Quality & Selection -Adjusted Measures			
	(M)	(S)	(M)	(S)	(M)	(S)	(M)	(S)	(M)	(S)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$1/(1-\nu)$	1.72***	1.24***	1.28***	1.15***	1.31***	1.38***	1.63***	1.31***	1.77***	1.56***
	(0.138)	(0.049)	(0.066)	(0.045)	(0.115)	(0.116)	(0.056)	(0.042)	(0.092)	(0.086)
Time Diff.(Δ)	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument	No	No	No	No	Yes	Yes	No	No	Yes	Yes
N	408	408	306	306	306	306	267	267	267	267

Table 4: Comparison of Elasticity Estimates: Different Specifications

Notes.–Sources: Census Bureau, 1990-2015. The table reports the elasticity of substitution between information technology (IT) and non-IT workers in two ways. Columns 1-6 estimate the elasticity according to the following equation

$$\Delta \ln\left(\omega_i^{IT} L_i^{IT}/\omega_i^{NIT} L_i^{NIT}\right) = \Delta_t \ln\left(\alpha_i/(1-\alpha_i)\right) + \nu_i \Delta_t \ln\left(L_i^{IT}/L_i^{NIT}\right) + \nu_i \Delta_t \left(\varepsilon_i^{IT}/\varepsilon_i^{NIT}\right)$$

where our outcome variable is the ratio of the wage bills and our right-hand-side variable is the ratio of effective labor services, which we estimate in log-levels and first-differences, using ordinary least squares and instrumental variables. We obtain quality-adjusted labor services premia by first estimating a group × year skill price obtained by regressing annual earnings on an indicator for working in an IT job, conditional on controls, and subsequently dividing total earnings by the skill price proxy and multiplying by an individual's total hours worked during the year. Controls in the skill price estimation include race, family size, marital status, gender, a quadratic in years of schooling, and age bin fixed effects (20-29, 30-39, 40-49, 50+). Controls in our estimation in the equilibrium condition include: group-specific average family size, share of males, race shares, a quadratic in schooling, and age bin shares. Columns 7-10 estimate the elasticity according to the following equation

$$\Delta \ln\left(\omega_i^{IT}/\omega_i^{NIT}\right) = \Delta_t \ln\left(\alpha_i/(1-\alpha_i)\right) + (\nu_i - 1)\Delta_t \ln\left(L_i^{IT}/L_i^{NIT}\right) + \nu_i \Delta_t \left(\varepsilon_i^{IT}/\varepsilon_i^{NIT}\right)$$

where the difference from columns 1-6 is that our skill prices and quality-adjusted measure of labor services are obtained from our selection model based on Adao (2016). Controls in all specifications are the same and our instruments include the IT employment premia growth between 1970 and 1980 and the share of IT workers in 1970. Observations are weighted by the number of workers observed in each cell in 1980.

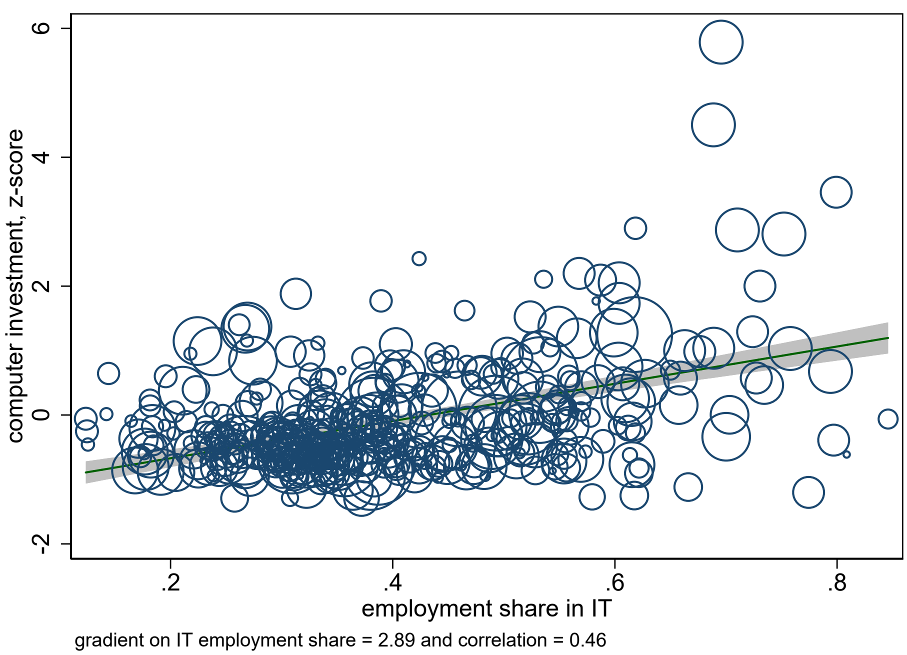


Figure 9: Comparison of Physical Computer Investment and Employment IT Shares

Notes.–Sources: Current Population Survey and Acemoglu et al. (2014), 1970-2009. The figure plots the share of computer investment relative to total (using Acemoglu et al. (2014)) with the employment share of IT (using the baseline O*NET measure) at a three-digit industry level of aggregation. Observations are weighted by the number of individuals in each three-digit industry cell obtained from the CPS. The figure shows a positive correlation between computer investment and IT workers in the cross-section of industries.

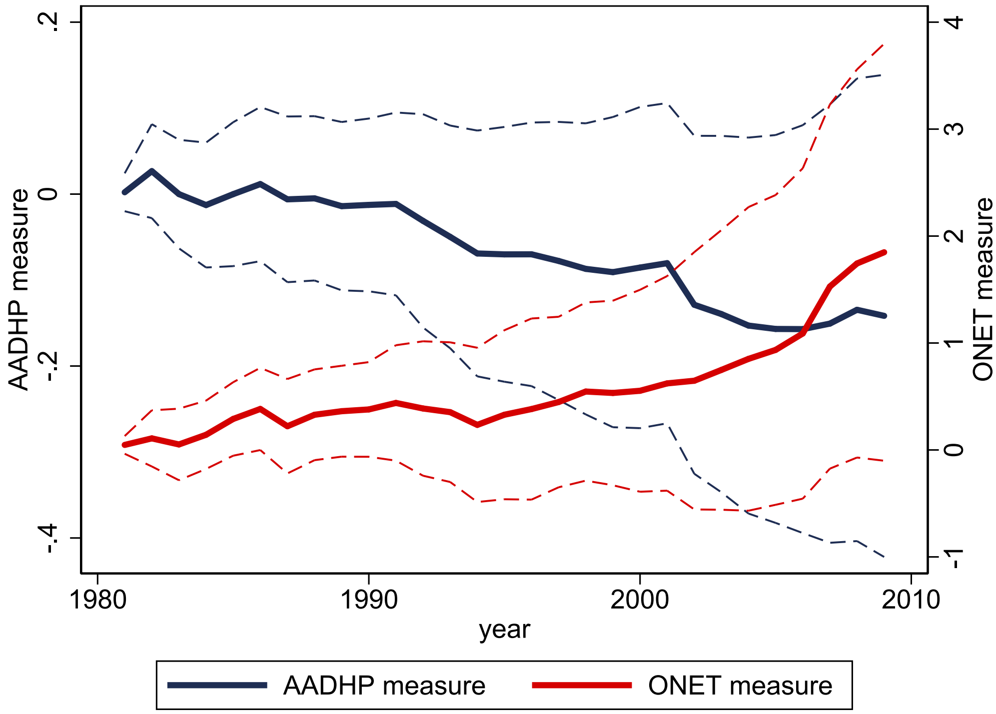


Figure 10: IT Intensity and Real Value of Shipments, AADHP and O*NET Measure

Notes.–Sources: Current Population Survey (1980-2009) and Acemoglu et al. (2014). The figure plots the coefficient on the interaction between the IT intensity measure and year dummies (normalized to 1980 as zero) from regressions of logged real value added and real payroll expenditures on year dummies, three-digit census industry dummies, and the interaction between IT and year dummies. Standard errors are clustered at the industry level and observations are weighted by the industry employment share.

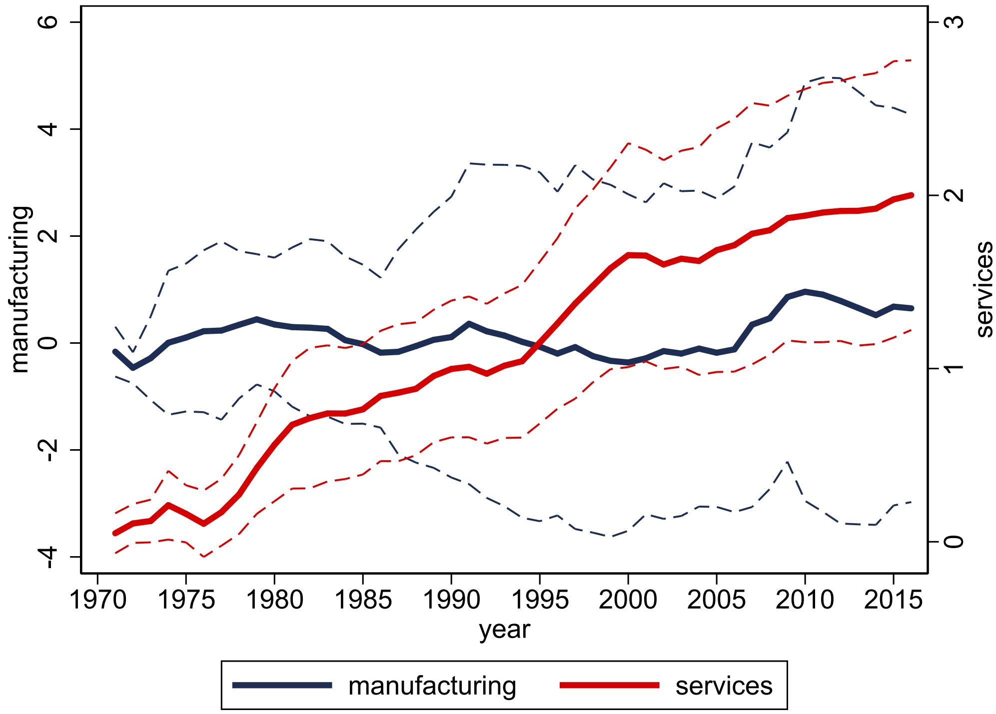


Figure 11: IT Intensity and Real Output, Manufacturing and Services
Notes.–Sources: Current Population Survey and Bureau of Economic Analysis, 1970-2016. Consider the following regression: $y_{jt} = \alpha + \lambda_t + \sum_t \delta^t (IT_j \times \lambda_t) + \eta_j + \epsilon_{jt}$. The figure plots the coefficients associated with the interaction, δ^t, separately for manufacturing and services sectors. The outcome variable is real gross output, which is deflated using 2009 price indices. IT denotes the employment share of IT at a three-digit NAICS level of aggregation between 1970 and 2015. Standard errors are clustered at the industry level and observations are weighted by average employment between 1998-2015.

A Online Appendix (Not for Print)

A.1 Supplement to Data Measurement

A.1.1 Index Measures of IT Intensity

We use a combination of input data from O*NET on the underlying tasks, knowledge, and skills workers use at a six-digit level of occupational heterogeneity. While we detail the different indices below, we have also experimented with different clustering algorithms besides our baseline approach of averaging across each of these indices.

Computers and Electronics (knowledge): Knowledge of circuit boards, processors, chips, electronic equipment, and computer hardware and software, including applications and programming Question: What level of COMPUTERS AND ELECTRONICS is needed to perform your current job? 1 (lowest level): Operate a VCR to watch a pre -recorded training tape 4 (intermediate): Use a word processor 6 (high): Create a program to scan computer disks for viruses.

Interacting with computers (work activity): Using computers and computer systems (including hardware and software) to program, write software, set up functions, enter data, or process information. Question: What level of WORKING WITH COMPUTERS is needed to perform your current job? 1 (lowest level): Enter employee information into a computer database 4 (intermediate): Write software for keeping track of parts in inventory 6 (high): Set up a new computer system for a large multinational company.

Programming (skills): Writing computer programs for various purposes Question: What level of PROGRAMMING is needed to perform your current job? 1 (lowest level): Write a program in BASIC to sort objects in a database 4 (intermediate): Write a statistical analysis program to analyze demographic data 6 (high): Write expert system programs to analyze ground radar geological data for probable existence of mineral deposits.

System (skill): Evaluation Identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system Question: What level of SYSTEMS EVALUATION is needed to perform your current job? 1 (lowest level): Determine why a coworker has been overly optimistic about how long it would take to complete a task 4 (intermediate): Identify the major reasons why a client might be unhappy with a product 6 (high): Evaluate the long-term performance problem of a new computer system.

Quality control analysis (skill): Conducting tests and inspections of products, services, or processes to evaluate quality or performance Question: What level of QUALITY CONTROL ANALYSIS is needed to perform your current job? 1 (lowest level): Inspect a draft memorandum for clerical errors 4 (intermediate): Measure new part requisitions for tolerance to specifications 6 (high): Develop procedures to test a prototype of a new computer system.

Operations analysis (skill): Analyzing needs and product requirements to create a design Question: What level of OPERATIONS ANALYSIS is needed to perform your current job? 1 (lowest level): Select a photocopy machine for an office 4 (intermediate): Suggest changes in software to make a system more user friendly 6 (high): Identify the control system needed for a new process production plant.

Updating and using Relevant knowledge (work activity): Keeping up-to-date technically and applying new knowledge to your job Question: What level of UPDATING AND USING RELEVANT KNOWLEDGE is needed to perform your current job? 1 (lowest level): Keep up with price changes in a small retail store 4 (intermediate): Keep current on changes in maintenance procedures for repairing sports cars 6 (high): Learn information related to a complex and rapidly changing technology.

Technology design (skill): Generating or adapting equipment and technology to serve user needs. Question: What level of TECHNOLOGY DESIGN is needed to perform your current job? 1 (lowest level): Adjust exercise equipment for use by a customer 4 (intermediate): Redesign the handle on a hand tool for easier gripping 6 (high): Create new technology for producing industrial diamonds.

Analyzing Data or Information (work activity): Identifying the underlying principles, reasons, or facts of information by breaking down information or data into separate parts. Question: What level of ANALYZING DATA OR INFORMATION is needed to perform your current job? 1 (lowest level): Determine the location of a lost order 4: (intermediate): Determine the interest cost to finance a new building 6: (high): Analyze the cost of medical care services for all hospitals in the country.

Processing Information (work activity): Compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data. Question: 1 (lowest level): Tabulate the costs of parcel deliveries 4: (intermediate): Calculate the adjustments for insurance claims 6: (high): Compile data for a complex scientific report.

Engineering and technology (knowledge): Knowledge of the practical application of engineering science and technology. This includes applying principles, techniques, procedures, and equipment to the design and production of various goods and services. Question: What level of knowledge of ENGINEERING AND TECHNOLOGY is needed to perform your current job? 1 (lowest level): Install a door lock 4: (intermediate): Design a more stable grocery cart 6: (high): Plan for the impact of weather in designing a bridge.

Management of Material Resources: Obtaining and seeing to the appropriate use of equipment, facilities, and materials needed to do certain work. Question: What level of MANAGEMENT OF MATERIAL RESOURCES is needed to perform your current job? 1 (lowest level): Rent a meeting room for a management meeting 4: (intermediate): Evaluate an annual uniform service contract for delivery drivers 6: (high): Determine the computer system needs of a large corporation and monitor use of the equipment.

We take the sum of these constructed indices and produce a standardized z-score with a mean of zero and standard deviation of one. We classify jobs as high (low) IT if the z-score is above (below) zero.

A.1.2 Background Comparisons of Indices

Here we discuss comparisons of our IT index to alternative measures. Namely, we examine how results change when we use a close analogue that classifies jobs as high IT if they are above the median z-score of IT intensity. Figure 12 plots the earnings and employment premia across the manufacturing and services sectors under both the baseline (mean) and alternative (median) score rule. These premia are defined as the logged earnings and employment difference between high and low IT jobs, respectively. Panels A and C illustrate that the earnings premia are very similar in both levels and trends—the correlation between the two is 0.83. For example, where the definition based on the mean implies an initial 1970 logged earnings difference around 0.50 in manufacturing, the definition based on the median implies an earnings difference of roughly 0.39.

Panels B and D show that the employment premia have similar trends. The differences in levels should not be surprising in light of the fact that the definition based on the median contains a larger number of observations since it splits occupations evenly into high and low IT, whereas the definition based on the mean includes roughly 7% fewer observations in the high IT category based on the distribution of IT intensity across jobs.

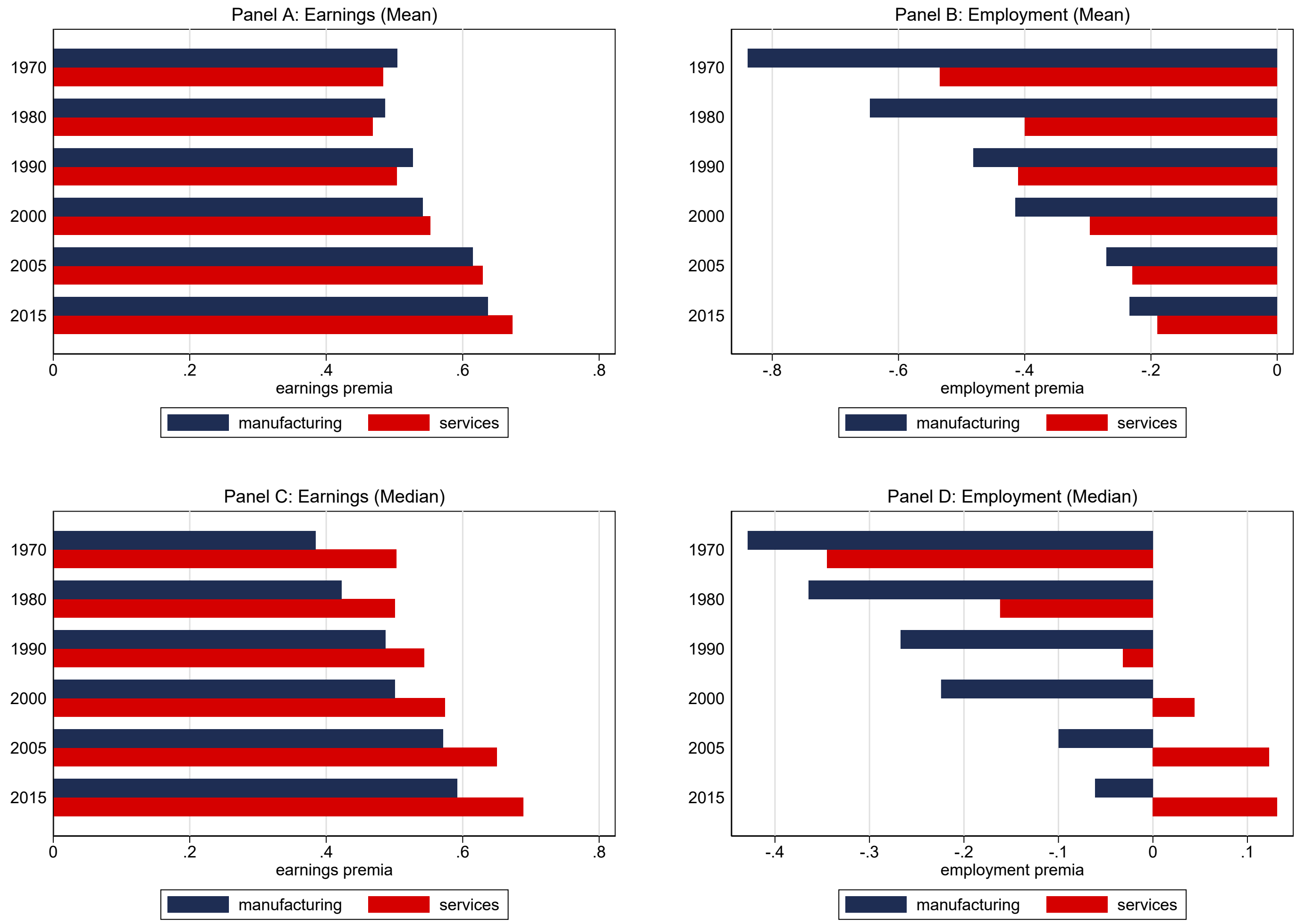


Figure 12: Comparison of Mean and Median Based Information Technology Labels

Notes.–Sources: Census Bureau and O*NET, 1970-2015. Panels A and C plot the logged (average) annual earnings among IT workers net of logged annual earnings among non-IT workers (weighted by Census sample weights to produce the average) in the manufacturing and services sector using the two different definitions of IT jobs. Earnings is deflated using the 2010 personal consumption expenditures index and the sample is restricted to workers earning over \$5,000/year, \$2/hour, and working 500 hours/year. Panels B and D plot the logged total number of IT workers net of the logged total number of non-IT workers using the two different definitions of IT jobs. In the "mean" based definition, IT workers are classified as those in an occupation with an IT intensity score above the mean (z-score of zero) in the distribution of five-digit occupations, whereas, in the "median" based definition, IT workers are classified as those in an occupation with an IT intensity score above the median z-score in the distribution of five-digit occupations.

We also examine an alternative measure of IT intensity among workers, proposed by the Census (Beckhusen, 2016). Initially introduced in 1970, this classification designated only three occupational groups as IT jobs, and by 2010 it included 12 occupational groups. Figure 13 plots both the employment count of IT workers and the employment share of IT jobs in the overall economy since 1970. This measure captures jobs that overwhelmingly use computers (e.g., database administrators and computer programmers), while omitting a wide array of occupations that regularly interact with other types of information technology. We, therefore, conclude that our approach is preferable as the O*NET provides a systematic way of surveying all occupations and their relative interaction and dependence on IT related skills, task, and knowledge.

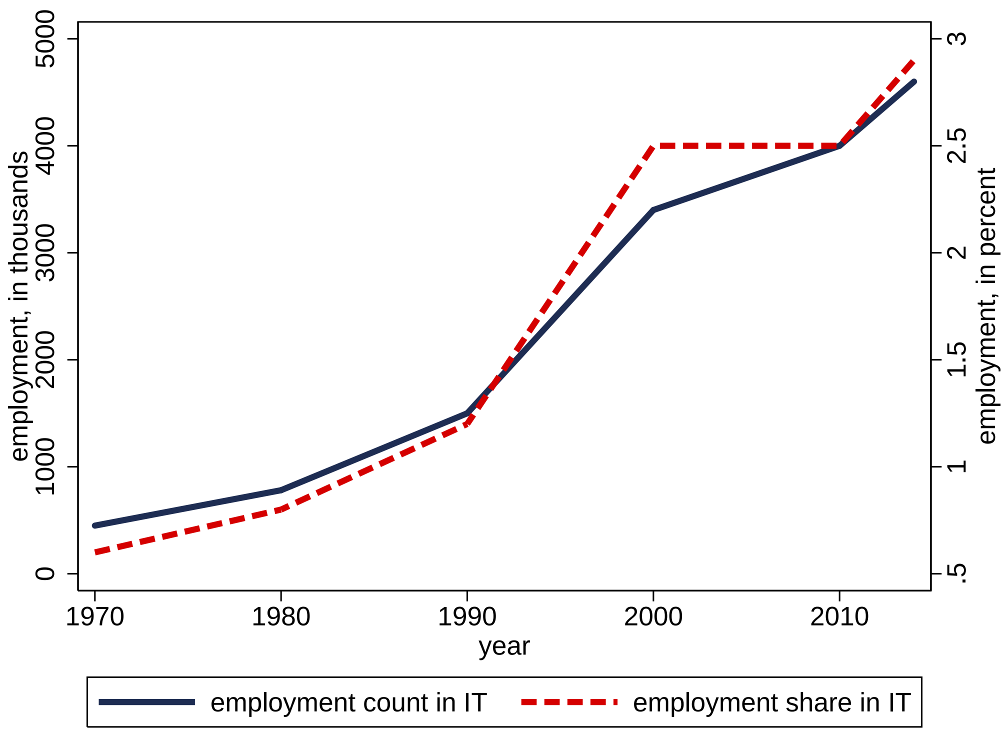


Figure 13: Information Technology Workers, 1970-2014

Notes.–Sources: See Beckhusen (2016) for calculations. Occupation codes were harmonized to the 2010 census classification.

A.2 Supplement to Descriptive Statistics

The main text presents regressions of various outcome measures (e.g., logged hourly wage) on IT intensity using Census micro-data. While these coefficients characterize the conditional mean, we now explore the distribution of these outcomes in occupations with high and low IT using six-digit occupational variation from the Occupation and Employment Statistics (OES), displayed in Figure 14. The benefit of these data is that they are comprehensive and contain the most granular variation available that can match with our IT intensity index. We find that there is a remarkable difference in the distribution of hourly wages between jobs with more versus less IT intensity. The difference in employment and inequality between the two are a little more subtle, although high IT jobs tend to be larger in both respects.

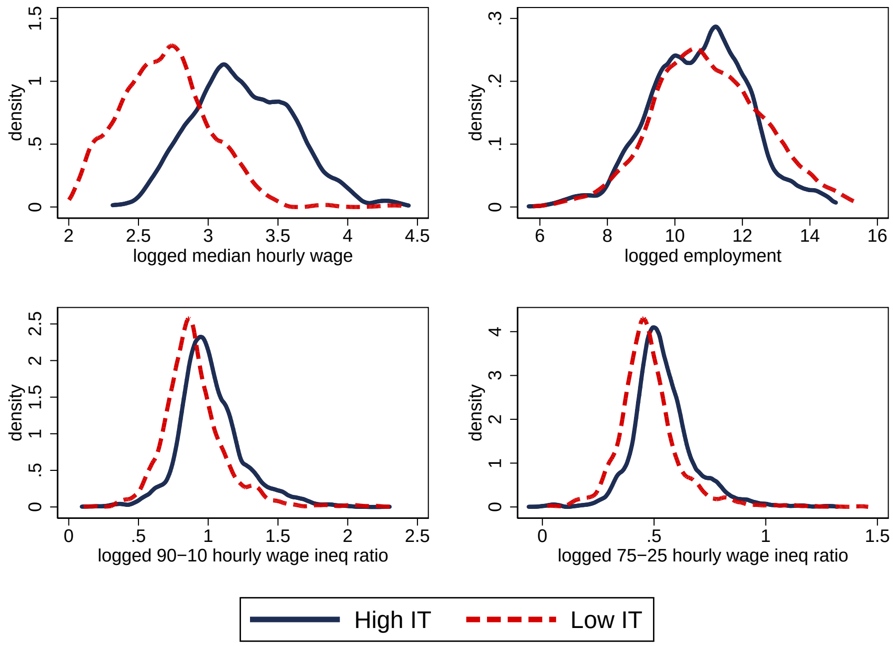


Figure 14: Distribution of Hourly Wages, Employment, and Inequality, by IT Intensity

Notes.–Sources: Bureau of Labor Statistics Occupation and Employment Statistics (national) tables, 2000-2015. The figure plots the distribution of logged median hourly wages, employment, and hourly wage ratio inequality proxies for high and low IT-intensive occupation.

We subsequently examine the cross-sectional differences in earnings and employment premia by major industry. Figure 15 plots logged income in IT-intensive jobs net of income in non-IT-intensive jobs, together with the IT employment share, separately by major industry, using the Census micro-data for 1980 and 2013-2015. Whereas in some industries there is a very small premium (e.g., wholesale), in others there is a high premium. For example, in finance, insurance, and real estate, IT-intensive workers earn approximately 70% more than their counterparts. While both FIRE and manufacturing sectors have similar IT income premia, the employment share of IT workers is much greater in FIRE than it is in manufacturing (e.g., 40% versus 25%).

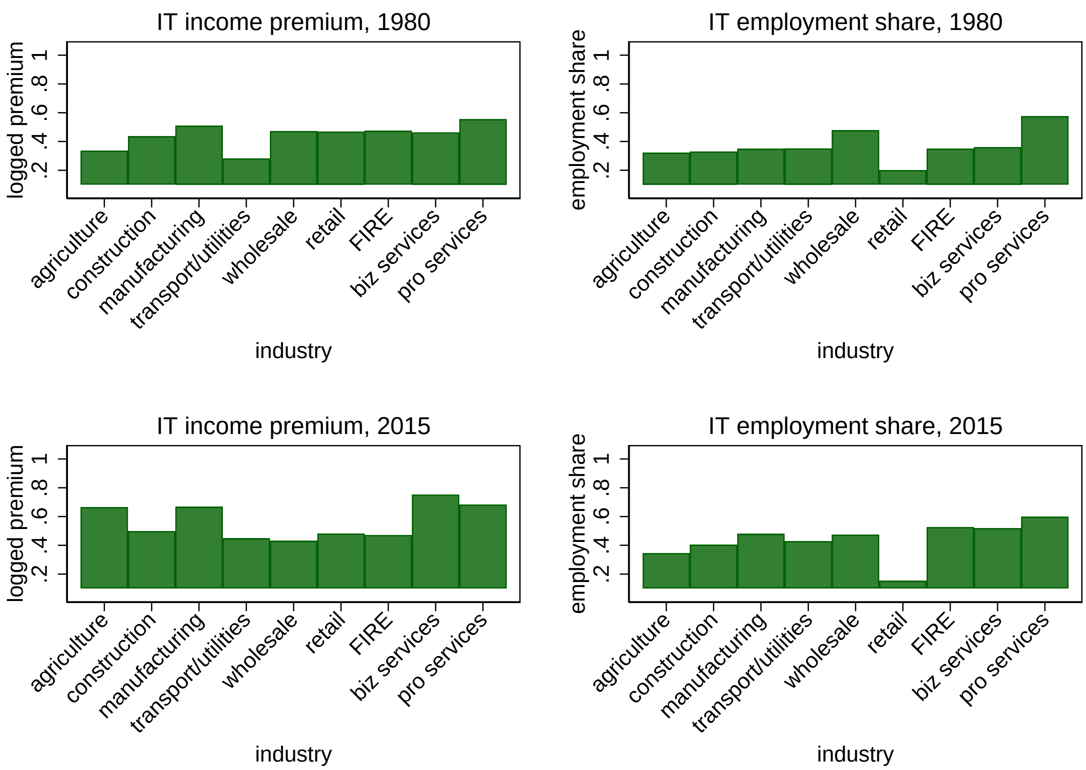


Figure 15: Earnings and Employment Premia, by Industry and Year Group

Notes.–Sources: Census Bureau and O*NET, 1980 and 2013-2015. The figures plot the logged IT earnings premium obtained by taking logged labor income in IT-intensive jobs net of logged labor income from non-IT-intensive jobs across, and the IT employment share, both across industries. Observations are weighted by the survey sample weights.

A.3 Supplement to the Descriptive Evidence and Facts

A.3.1 Robustness of Controlling for Skill Intensity

The main text presents evidence showing that hourly wages are significantly higher in jobs with high IT intensity, relative to low intensity. However, once concern is that IT intensity is simply correlated with other valuable skills through selection channels. We address this concern by exploiting cross-sectional variation through regressions of the form

$$y_{ot} = \alpha skill_o + \gamma_t IT_o + \psi_{o'} + \epsilon_{ot} \tag{7}$$

where y denotes our outcome variable of interest (logged employment, inequality, and hourly wages), $skill$ denotes a vector of occupation-specific skills, IT denotes the intensity of information technology, and $\psi_{o'}$ denotes fixed effects on four-digit occupation cells. We present two main sets of estimates for Equation 7: unconditional and conditional correlations. The conditional correlation estimates illustrate that the return to IT intensity is large, even after controlling for the skill content associated to different tasks. These fixed effect estimates also illustrate that the wage premium persists after controlling for non-random sorting of different workers into different occupations at a detailed four-digit level.

Table 5 documents these results. Beginning with the logged hourly wage as the outcome variable, the unconditional correlation estimate suggests that a standard deviation rise in information technology is associated with a large 0.45% rise in the median logged hourly wage. The estimate is still statistically and economically significant after introducing detailed measures of skill intensity and four-digit fixed effects. Importantly, IT intensity is approximately half as large in magnitude as the association between hourly wages and non-routine & cognitive skills, which suggests that IT intensity is not merely a proxy for high skilled occupations. We also find that, in the cross-section, a standard deviation increase in IT intensity is associated with a large 0.57% decline in occupational employment and a 0.12% rise in the 90-10 logged hourly wage difference.

While the employment and inequality differences between IT and non-IT jobs are stark in the cross-section, they become statistically insignificant once we control for four-digit occupation and year fixed effects. Only an hourly wage premium remains: a standard deviation rise in IT intensity is associated with a 0.13% rise in the hourly wage. Regressions that control for occupation and year fixed effects also contain the standard set of skill intensity measures from the skill-biased technical change literature (Acemoglu and Autor, 2011). Strikingly, a standard deviation rise in non-routine and cognitive skills is associated with almost as large of an increase in the hourly wage as a proportional increase in IT intensity—0.12 for non-routine & cognitive versus 0.13 for IT. We also find that an increase in non-routine & cognitive is associated with a statistically significant decline in employment and a statistically significant rise in inequality, consistent with early evidence from Autor et al. (2003).

A.3.2 Time Series Variation in IT Tasks Intensity

We provide a brief characterization of the time series heterogeneity in aggregate IT intensity by plotting the evolution of each of its sub-components between 2000 and 2016, weighted by each six-digit occupation's employment from the Occupational Employment Statistics (OES) data. Table 6 reports the mean and standard deviation of each input intensity over time. Across most of the categories, the intensity is growing and generally matches our intuition about the types of tasks

Table 5: IT Intensity and Task Content of Jobs

Dep. var. =	logged hourly wage		logged employment		logged 90-10 ratio	
	(1)	(2)	(3)	(4)	(5)	(6)
information technology	.45***	.13***	-.57***	.02	.12***	.00
	[.02]	[.02]	[.10]	[.09]	[.02]	[.01]
non-routine, cognitive		.12***		-.28***		.05***
		[.02]		[.11]		[.02]
non-routine, non-cognitive		-.02		.06		-.01
		[.02]		[.08]		[.02]
routine, cognitive		.04***		-.05		.01
		[.01]		[.06]		[.01]
routine, manual		-.02		-.14		-.09***
		[.02]		[.12]		[.03]
non-routine, manual		.03		.17		.04*
		[.03]		[.12]		[.02]
R-squared	.63	.92	.13	.64	.22	.73
Sample Size	6897	6897	7516	7516	6979	6979
Controls	No	Yes	No	Yes	No	Yes
4-dig Occupation FE	No	Yes	No	Yes	No	Yes
Year FE	No	Yes	No	Yes	No	Yes

Notes.–Sources: Occupation and Employment Statistics (BLS) and O*NET, 2004-2015. The table reports the coefficients associated with regressions of logged employment, median hourly wages, and the logged 90-10 hourly wage differential on a standardized measure of IT intensity and standardized measures of skills following the strategy in Acemoglu and Autor (2011). Standard errors are clustered at the six-digit occupation level.

that have become more common, such as processing information, analyzing data, and task related to engineering and technology. However, dispersion in the intensity is staying roughly constant across most categories.

We now provide another way of characterizing the heterogeneity more visually. In particular, looking at the mean IT intensity might confound heterogeneity in the distribution of IT intensity across occupations. While the main text provides some useful heuristics that display the average IT intensity over time, it confounds a significant amount of heterogeneity across 773 unique six-digit occupations. Figure 16, therefore, plots the distribution of each input to the aggregate index in both 2004-2006 and 2014-2016 time periods. While a few of the distributions are relatively time invariant (e.g., quality control and programming), many distributions exhibit interesting changes. For example, both the "computers & electronics" and "interaction with computers" categories grow in both mean and skewness.

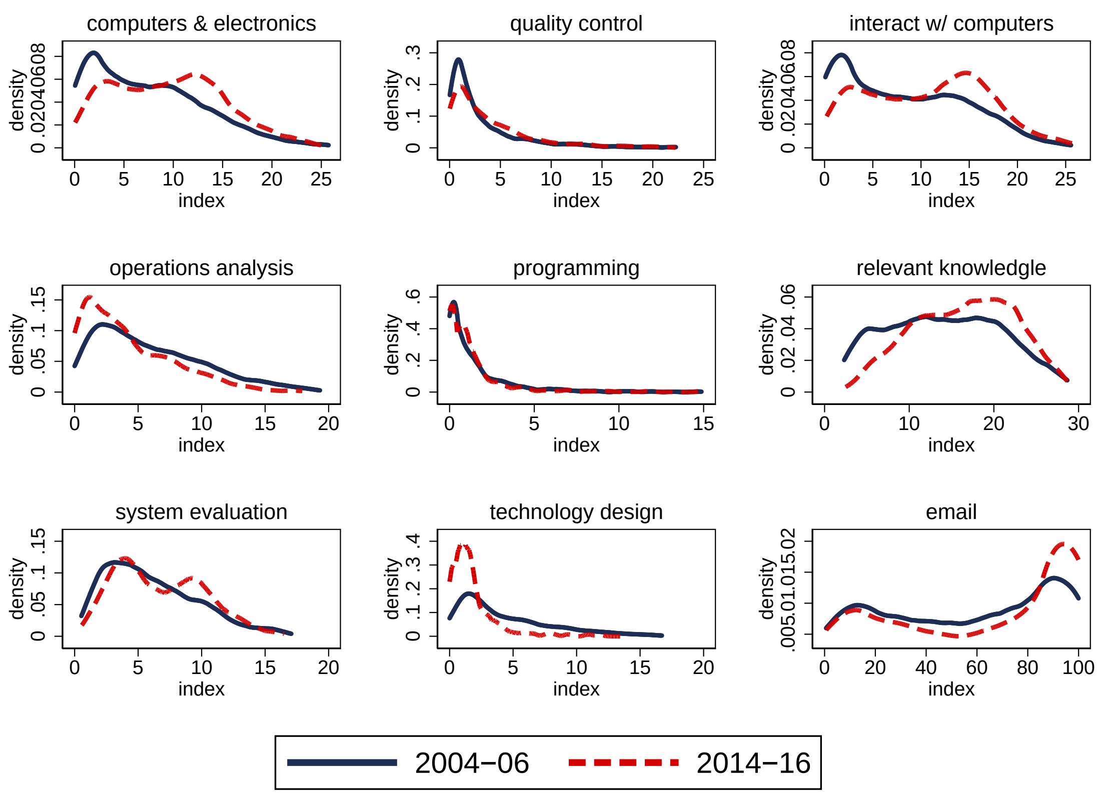


Figure 16: Distribution of Inputs to Information Technology Intensity

Notes.–Sources: O*NET. The figure plots the intensity of the nine inputs to the aggregate information technology (IT) index for 2004 and 2016. The annual measures are weighted across six-digit occupations using average employment from the Occupation Employment Statistics (OES).

Table 6: Time Series of Technology and Information Services Inputs

	2004-2005		2006-2007		2008-2009		2010-2011		2012-2013		2014-2015	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
computers and electronics	6.4	4.9	7.3	5.0	7.5	5.0	8.2	5.6	8.6	5.9	8.8	5.9
drafting	2.2	3.2	2.4	3.3	2.7	3.5	2.8	3.7	2.9	3.6	2.8	3.5
interacting with computers	8.3	6.1	8.5	6.0	8.9	6.0	9.6	6.1	10.1	6.2	10.4	6.2
operations analysis	4.6	4.0	5.8	4.2	4.8	3.9	3.5	3.6	3.4	3.6	3.2	3.4
programming	1.2	2.1	1.4	1.7	1.4	1.6	1.0	2.0	1.2	2.4	1.2	2.3
relevant knowledge	12.2	6.2	12.6	5.8	13.0	5.8	13.9	5.8	14.3	5.8	14.4	5.8
systems evaluation	5.3	3.8	5.1	3.6	5.3	3.7	5.7	3.6	6.0	3.6	6.0	3.6
technology and design	2.9	2.9	4.2	3.4	3.7	3.1	1.2	1.6	1.4	1.8	1.4	1.7
engineering and technology	2.8	4.5	3.2	4.6	3.4	5.0	3.7	5.0	3.7	4.9	3.7	4.8
managing resources	4.1	3.4	5.2	3.9	4.8	3.6	2.9	2.6	2.9	2.5	2.9	2.5
analyzing data	8.8	5.8	8.4	5.4	8.7	5.6	9.4	5.8	9.9	5.8	10.0	5.8
processing information	10.8	5.6	11.1	5.4	11.5	5.5	12.0	5.6	12.3	5.5	12.5	5.5
Observations	1381		1376		1389		1451		1520		1540	

Notes.–Sources: O*NET, 2000-2016. The table reports the means and standard deviations of each information technology sub-index at a six-digit level of occupation aggregation weighted by employment. Observations are weighted by employment.

A.3.3 Robustness: Labor Productivity Measures

The main text reports time series for real labor productivity, defined as value added per hour worked, using KLEMS data from 1950 to 2010. Sectors are defined as high IT if the weighted z-score of IT intensity, obtained from matched O*NET and Census micro-data for 2000, 2006-2008, and 2013-2015, is greater than zero. Our classification implies that NAICS22, NAICS324, NAICS325, NAICS333, NAICS335, NAICS336, NAICS339, NAICS51, NAICS52, and NAICS61 are high IT sectors, with the remaining sectors as low IT (NAICS11, NAICS21, NAICS23, NAICS311, NAICS314, NAICS321, NAICS322, NAICS326, NAICS327, NAICS331, NAICS42, NAICS44, NAICS48-49, NAICS531, NAICS532, NAICS62, NAICS71, NAICS72, and NAICS81).

One interesting result from our classification is the high value added per worker in the high IT manufacturing sector. A possible explanation behind these levels' differences is their extensive use of capital equipment, particularly since these are mostly primary metals sectors. Figure 17 addresses this concern by netting out capital compensation from value added, which produces a high IT services time series closer to high IT manufacturing. In fact, the growth in labor productivity under this approach is 700% for high IT services, whereas it is only 275% for high IT manufacturing (versus 506% and 327%, respectively, when using the baseline definition that does not net out capital compensation).

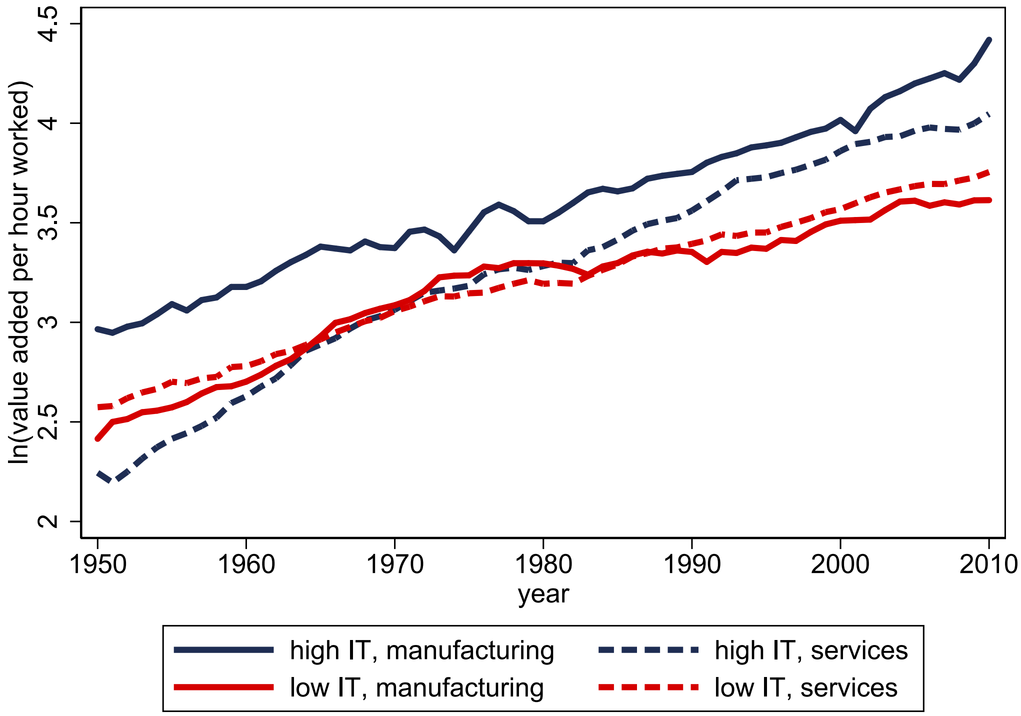


Figure 17: Value Added per Hour Worked in Manufacturing and Services, 1950-2010
Notes.–Sources: USA KLEMS, 1950-2010. The figure plots sectoral average annual value added *net of capital compensation* per worker in constant thousands of dollars (deflated using the 2009 GDP deflator from the St. Louis Fed) for high and low information technology manufacturing and services sectors. Sectors in each high/low IT and manufacturing/services category are averaged based on the long-run employment average between 1950 and 2010. Industries are classified as high information technology if the weighted average z-score of information technology intensity within three-digit NAICS codes is greater than zero. We construct these weighted average scores using a sample of Census micro-data for 2000, 2006-2008, and 2013-2015 matched with IT intensities at a five-digit level as implemented in the baseline definition. The figure shows that logged value added per hour worked has grown primarily in the high IT manufacturing and services sector with the bulk of the growth driven by high IT services.

As an additional point of comparison, we partition sectors in a different way. This allows us to visualize how different groupings of high and low technology sectors influence our understanding of heterogeneity within the services sectors. We use an alternative classification based on a "cherry-picking approach", in which we select sectors as high tech based on whether the sector sounds like one that makes intensive use of technology. In particular, we define high tech services sectors as publishing / software, motion picture / sound recording, broadcasting / telecommunications, information / data processing, federal reserve banks / credit intermediation, securities and commodity contracts, insurance carriers, funds / trusts, real estate, rental / leasing services, legal services, computer systems design / related, whereas low tech services sectors are wholesale trade, retail trade, social assistance, performing arts, amusements gambling / recreation, accommodation, food services / drinking, and other services.

We use value added in constant dollar of labor compensation as a measure of labor productivity

. We turn to this definition from the 2017 release of the KLEMS data because it holds constant potential differences in prices at a sectoral level, but it does not contain a measure of employment or total hours worked, which we would otherwise use as a closer point of comparison with our baseline. Figure 18 plots the resulting time series for manufacturing (pooled) and both high and low tech services. We find that the bulk of the increase in labor productivity is driven by the high tech services sector. While our selection criteria for high tech services is admittedly cherry-picked, these results show that our baseline approach provides, if anything, a conservative measure of value added in high tech services.

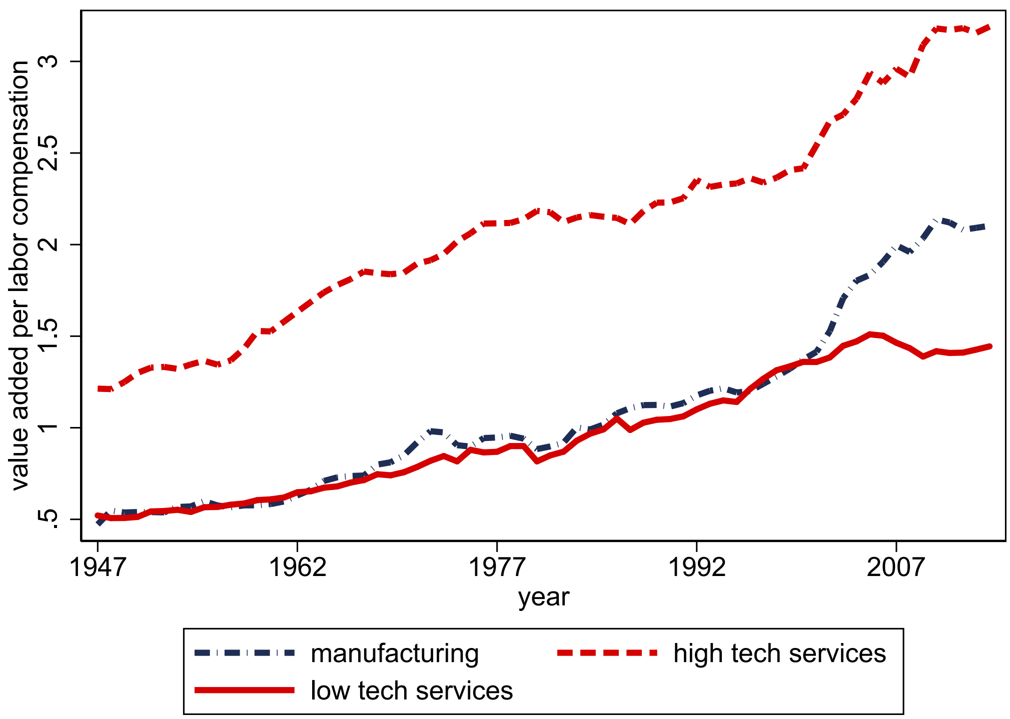


Figure 18: Value Added / Employee Compensation in Manufacturing and Services, 1947-2014
Notes.–Sources: USA KLEMS, 1947-2014. The figure plots sectoral average annual value added per labor (employee) compensation in constant thousands of dollars for manufacturing and services sectors. Sectors in each category are averaged based on the long-run employment average between 1947 and 2014. High skilled services sectors cover information and financial sub-sectors, including: publishing industries (including software), motion picture and sound recording, broadcasting and telecommunications, information and data processing, federal reserve banks credit intermediaries, securities commodity contracts, insurance carriers, funds, trusts and other financial vehicles, real estate, rental and leasing, legal, and computer systems design. Low skilled services sectors include wholesale / retail trade, social assistance, and other services (typically repair).

A.3.4 Robustness: Earnings and Employment Premia

We begin by showing further robustness of the earnings and employment premia to alternative refinements of the definition, namely the hourly wage and total hours (across all workers) premia. Panel A in Figure 19 shows a similar result as the main text that the hourly wage premium grew nearly identically in both manufacturing and services sectors from 1980-2013. The premium is also quite quantitatively close to the earnings premium, which suggests that hours differences play only a small role. Panel B tells a similar story. The observed total hours worked premium is similar as the employment premium in the main results, which suggests that the intensive margin differences in hours worked per worker in IT and non-IT jobs are fairly similar. There is a marginally larger quantitative narrowing of the IT total hours premium, relative to the employment premium, which follows from the fact that IT workers spend roughly 131 more hours worked per year (relative to their counterparts).[25]

[25] The difference is larger in the manufacturing sector with 143 hours worked more among IT workers versus only 125 hours worked more in the services sector. The estimate is conditional on family size, race, gender, marital status, and schooling.

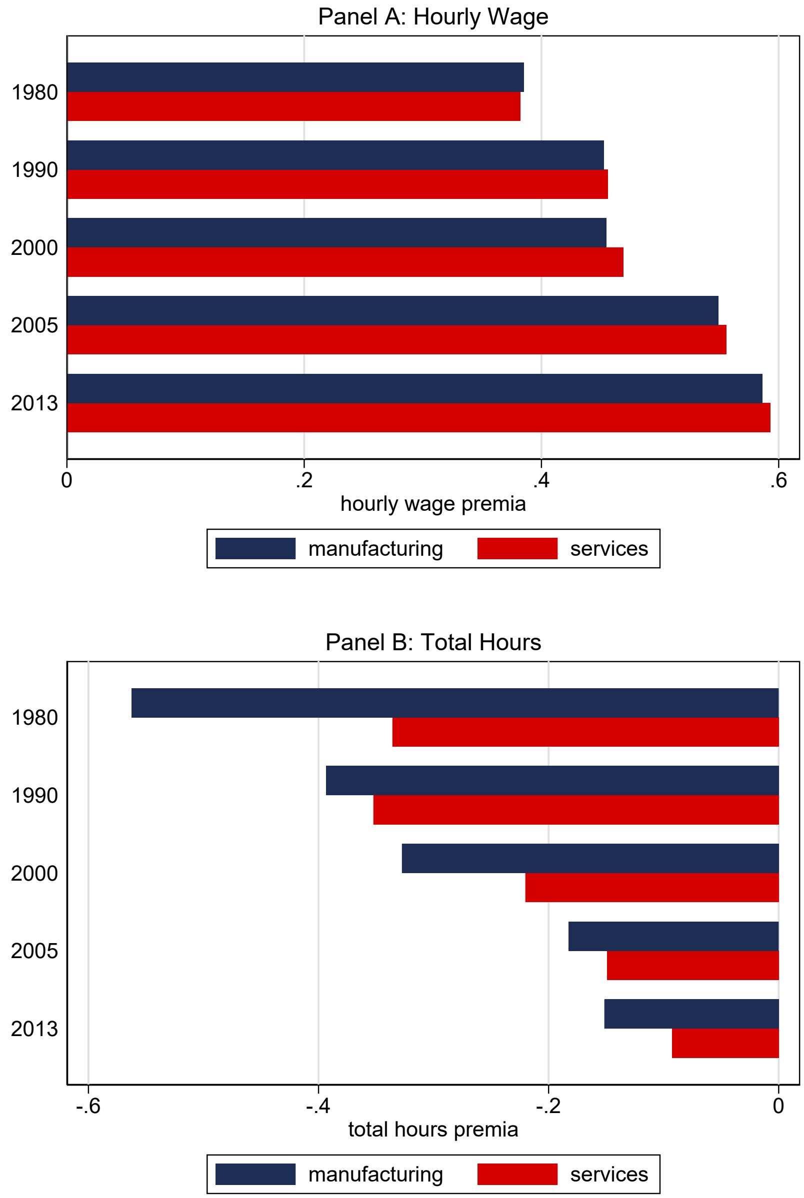


Figure 19: Hourly Wage and Total Hours Premia to Information Technology, 1980-2015
Notes.–Sources: Census Bureau and O*NET, 1980-2015. The figure plots the hourly wage and total hours (annual hours × number employees) in high technology and information services jobs between 1980 and 2015. Earnings is deflated using the 2010 personal consumption expenditures index. We start with 1980 because hours worked is not available in 1970.

We now turn to producing similar patterns on employment and earnings premia when using the annual CPS. The primary downside is that the estimates are more noisy due to sampling variability in the workers who are surveyed at the three-digit occupation level. Nonetheless, Figures 20 and 21 plot these premia. The crucial observation in Figure 20 is that, while the employment share in manufacturing is declining (thinned dotted blue line), the employment share of IT jobs in the manufacturing sector is rapidly rising (thick solid blue line). The rise of IT jobs accounts for nearly all of the growth in the services sector.

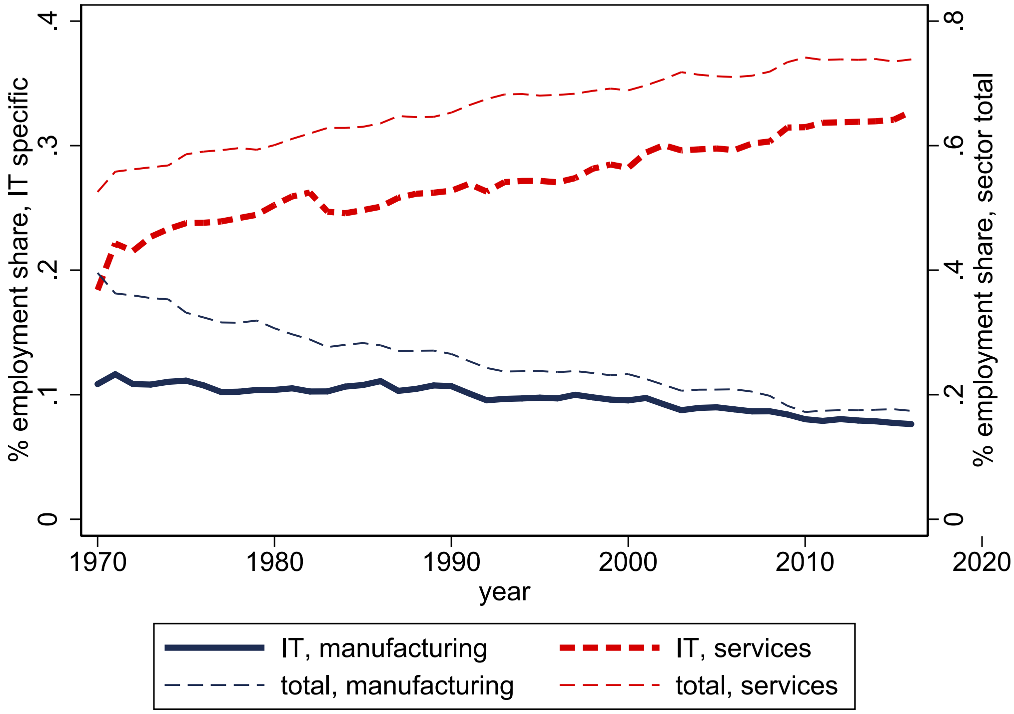


Figure 20: Employment Shares in Technology and Information Services—Sectoral and Aggregate, CPS Robustness

Notes.–Sources: Current Population Survey (CPS) ASEC, 1970-2015. The figure plots several employment shares. The left y-axis plots the employment share of manufacturing and services information technology workers ("IT v. non-IT") relative to their sectoral total. The right y-axis plots the employment share of manufacturing and services overall ("total"). As such, the level of the right y-axis will be larger than the level of the left y-axis because it includes non-IT workers as part of the share. The purpose of providing both sets of trends is to illustrate the IT share in manufacturing and services in light of the overall trends in these two sectors.

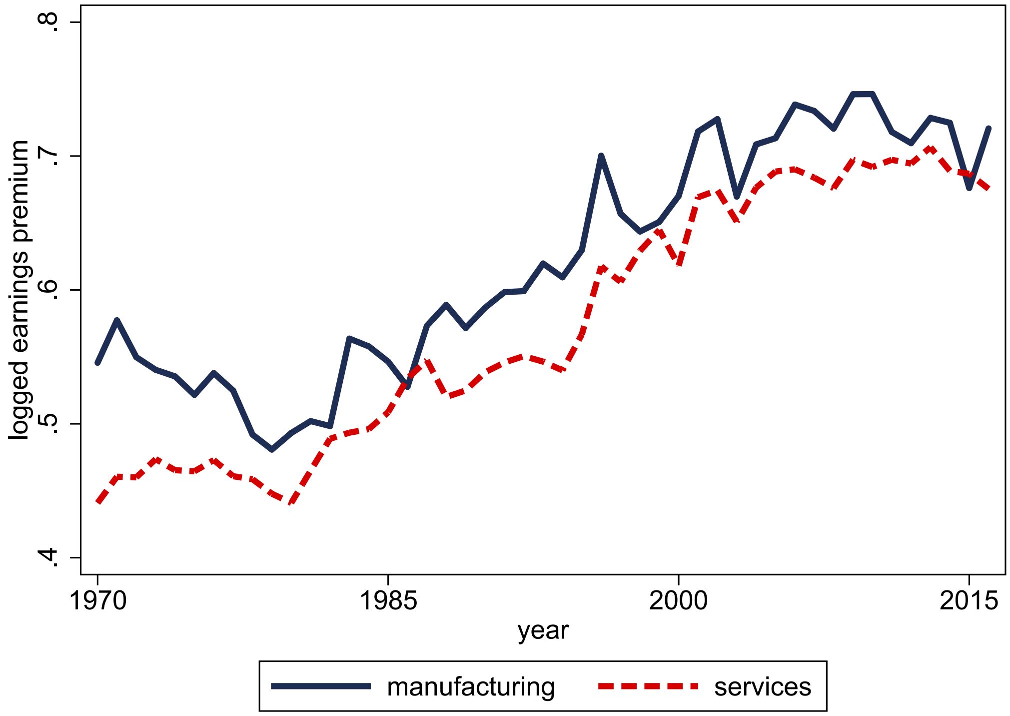


Figure 21: Earnings in Technology and Information Services, CPS Robustness

Notes.–Sources: Current Population Survey (CPS) ASEC, 1970-2015. The figure plots earnings premia in manufacturing and services sectors where the technology and information services (IT) intensity is based on being above the median score.

Figure 22 also examines heterogeneity in the earnings premia using two different classification strategies: K-medians and K-means with $K = 2$. K-medians produces an almost identical result as the regular medians-based approach, which is not surprising since our classification is based on the single IT score. Using K-means tends to classify more occupations as IT-intensive, which is why the earnings premia is marginally smaller—but clearly not by much by any stretch of the imagination.

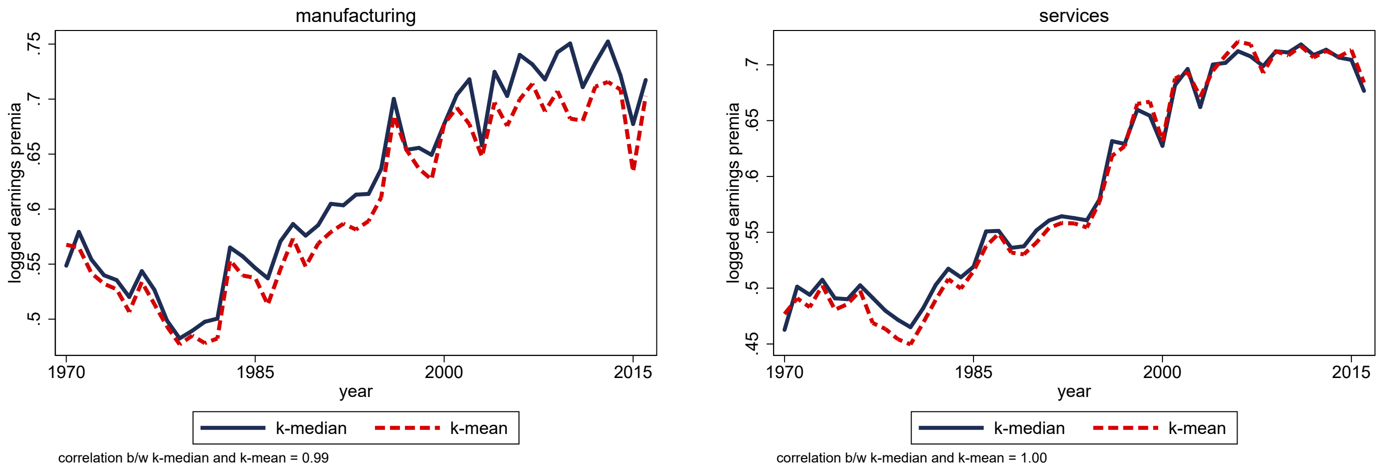


Figure 22: Comparing CPS Earnings Premia with K-medians and K-means

Notes.–Sources: Current Population Survey (CPS) ASEC, 1970-2015. The figure plots earnings premia in manufacturing and services sectors where the technology and information services (IT) intensity is based on two approaches: using K-medians and K-means estimators for $K = 2$.

While the two series produce qualitatively similar series (e.g., the correlation between the earnings and employment premia in the two are 0.88 and 0.86, respectively), there are at least two reasons the OES data generates quantitatively different series. The first is that, since the CPS does not aim to be representative of every occupation at a detailed five-digit level, it may overstate one type of worker over another. These concerns are potentially amplified by the presence of occupational misclassification, which has been documented by Kambourov and Manovskii (2013). The second is that the CPS data only contains a five-digit occupation classification, whereas the OES data contains a six-digit classification. To the extent there is some detailed within-occupation reallocation, differences can emerge.

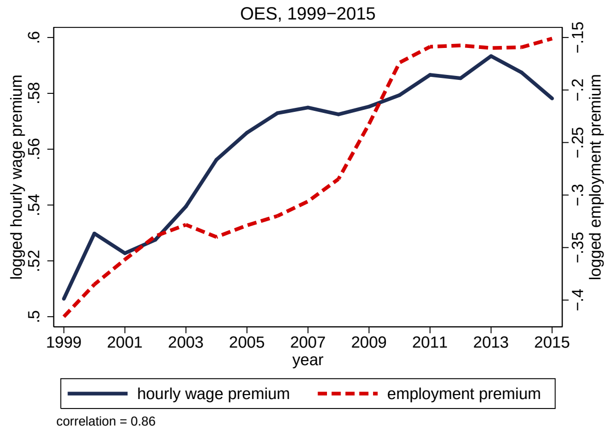


Figure 23: Earnings and Employment in Information Technology Jobs, National

Notes.–Sources: Bureau of Labor Statistics Occupational Employment Statistics, 1999-2015. The figure plots the logged earnings premium between information technology and non-IT workers (deflated using the 2010 personal consumption expenditures index) and the logged employment premium. The employment premium is computed by summing across all IT and non-IT workers at a six-digit occupation level.

Given the large rise in both the earnings and employment IT premia, we also examine how it interacts with other documented premia in the labor market, in particular the returns to tenure (we later examine the interaction with the returns to education). Using the CPS supplement on job tenure between 1996 and 2014, we are able to non-parametrically characterize the IT premium across the tenure distribution. To do this, we regress logged hourly wages on a vector of controlling covariates (age, number of children, race, gender, marital status, education), subsequently averaging across the residualized earnings measures for each tenure bin. We include these controls to mitigate the potential effects of the composition of the labor force throughout the tenure distribution. We separately plot these returns for the 1996-2002 and 2010-2014 time periods to understand the extent to which these returns might have shifted over time.

Figure 24 documents these results. First, and not surprisingly, the IT premium has grown over the past decade and it has shifted the earnings premium up across the entire tenure distribution. For example, between 1996-2002, IT workers with zero years of tenure would earn 30% more than their non-IT counterparts, whereas between 2010-2014 they would earn 37% more. Looking at the top of the tenure distribution, however, IT workers with 16-20 years of tenure earned roughly 24% more than their counterparts between 1996-2002, but between 2010-2014 the premium grew by only one to two percentage points.

Second, and more importantly, the IT premium is declining in tenure. While it is possible that technology companies—which employ a large share of IT workers—simply tend to have lower average tenure due to something embedded within their underlying production function, we provide evidence that the declining premium is a natural result of a career ladder where skilled workers begin in an IT-intensive job and then progress to a managerial job. Such examples are commonplace in many technology hubs, like Silicon Valley or San Francisco, where a skilled worker may begin as a data scientist or consulting out of undergraduate and then transition towards a senior managerial role after roughly a decade. The main reason for this is that, as individuals progress in the career ladder, managerial skills become increasingly important and outpace the importance of IT skills. In particular, while it is true that managers will leverage IT to broaden their span of control, their comparative advantage in management begins to outweigh their absolute advantage.

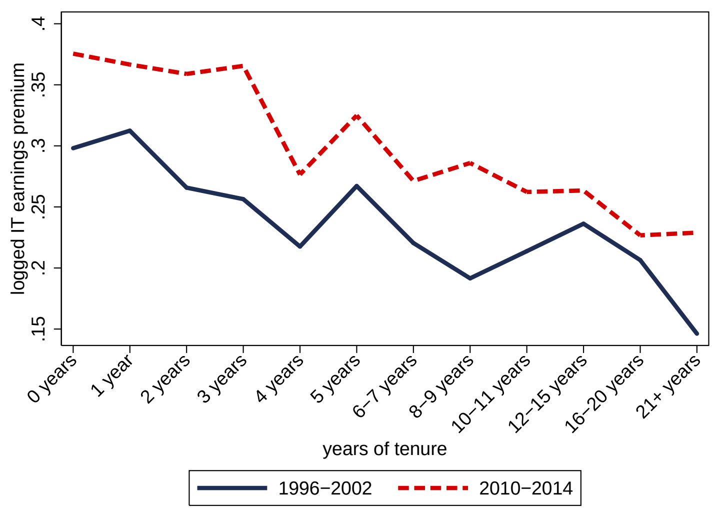


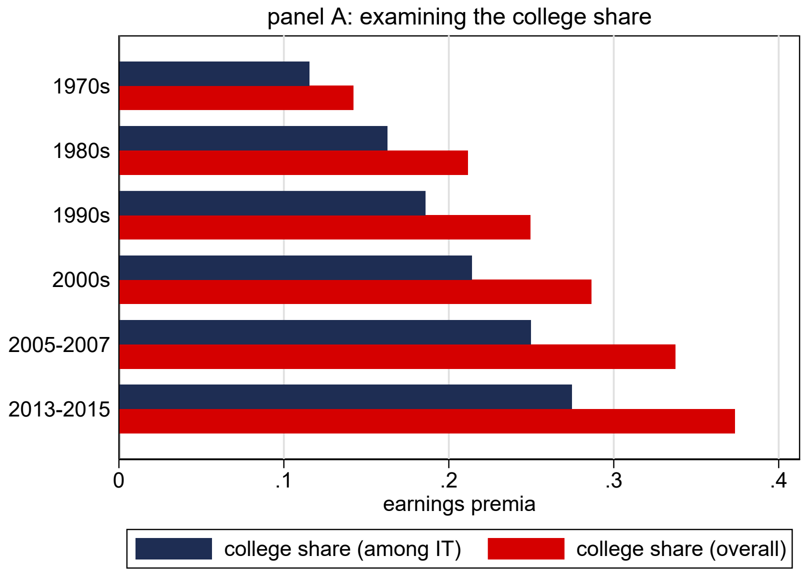
Figure 24: Hourly Wage Premium, by Employee Tenure

Notes.–Sources: Current Population Survey Job Tenure Supplement. The figure plots the residualized logged earnings premium between IT and non-IT-intensive jobs (identified using the O*NET measurements of tasks) across the distribution of tenure using the CPS tenure supplement covering 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, and 2014. Logged earnings are residualized by regressing them against a vector of controls, including age, education, race, gender, marital status, and number of children, taking the residual as the residualized logged earnings measure. Observations are weighted by the job tenure survey sample weights.

A.3.5 The Declining Premium for C/NR Tasks and College

The main text presents results about the earnings premium between IT and non-IT jobs restricted to the set of college degree workers, and separately for Cognitive and Non-Routine (C/NR) jobs. In both cases, the growth rate from 2000 and 2015 is greater within-group than it is across groups—that is, restricted to the set of college degree workers, the growth in the IT premium is greater than the growth in the college premium across all in the labor force. We now present complementary plots that characterize the employment premia between these jobs.

Beginning with Panel A in Figure 25, it is remarkable how much of the employment share of college workers is accounted for by high IT workers. However, there is important longitudinal variation. In the 1970s, these IT jobs accounted for almost all of the share of college degree workers; roughly 15% had a college degree and 12% were in IT jobs. However, by 2015, nearly 37% of the labor force has a college degree and 27% are in IT jobs. The gap between the college share and the "college + IT" share reflects the surge in growth of universities and expanded access to a four-year college degree. Turning towards Panel B in Figure 25, we see that there is not a substantial difference between the IT + C/NR and C/NR shares—that is, an even smaller difference in comparison with the college share. For example, in 1970, roughly 27% of jobs were classified as high C/NR and nearly 26% of those jobs were also classified as high IT. By 2015, the share of high C/NR jobs grew to 45% and the share of "IT + C/NR" jobs grew to nearly 40%.



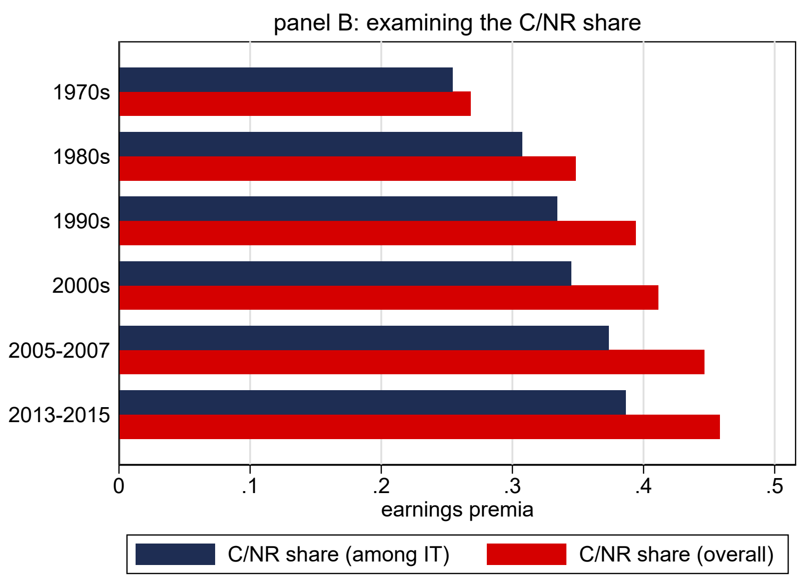


Figure 25: Examining Heterogeneity in the Demand for Skills in High and Low IT Jobs (Employment Shares)

Notes.–Sources: Census Bureau and O*NET, 1970-2015. Panel A plots the employment share of college and high IT workers relative to the total labor force, together with the overall college share. Panel B plots the employment share of high cognitive and non-routine and high IT occupations relative to the total labor force, together with the overall share of high C/NR jobs.

We examine the evolution of employment shares across the four permutations of IT/non-IT and college/non-college jobs in Figure 26. Importantly, the share of IT and college degree workers has grown from approximately 20% of the labor force to nearly 32%, and it dwarfs the marginal rise in non-IT and college degree workers from 2% to roughly 5% over the past 40 years. Symmetrically, the share of IT and non-college degree workers declined from 32% to 22%, which also dwarfs the decline in non-IT and non-college degree workers from 45% to 42%.

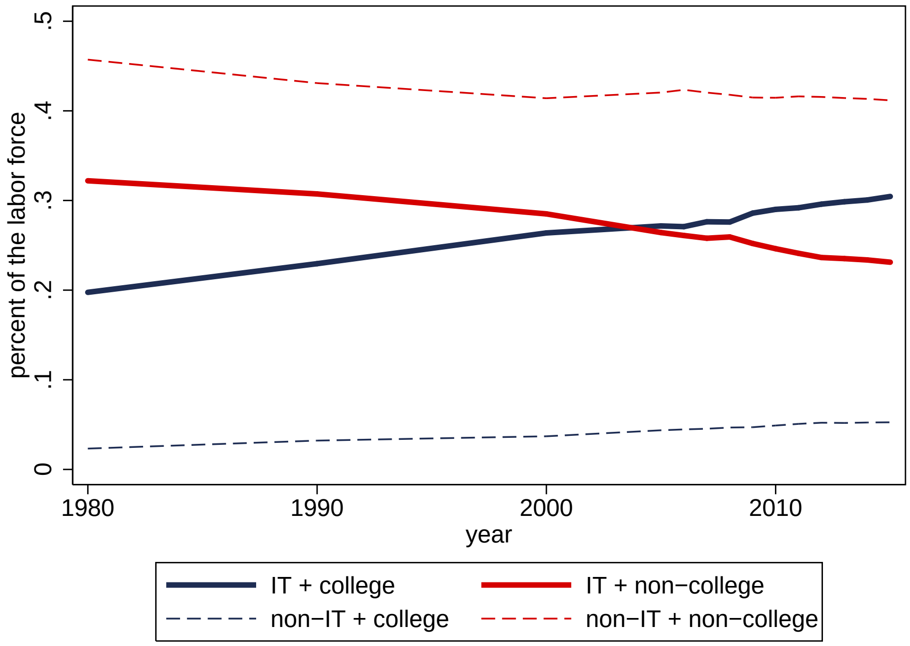


Figure 26: Employment Shares in Information Technology Jobs, by College Attainment
Notes.–Sources: Census Bureau, 1980-2015. The figure plots the share of full-time workers across the permutations of IT versus non-IT-intensive jobs and college and non-college degree workers.

A.3.6 IT Premia using the Census Classification

In the body of the paper we present a comparison of earnings and hours premia obtained using either our IT intensity measure or the industry-based measure from Hecker (2005). While these measures capture different features of the data—in part because Hecker (2005) focuses on industries, whereas we focus on jobs—we now turn towards an additional comparison of information technology jobs discussed by Beckhusen (2016). The Census began identifying IT jobs in 1970, capturing roughly 0.6 percent of the labor market. Beckhusen (2016) discusses how the measure has expanded since then and grown to 2.9% of the population. These workers represent roughly 16% of our Census sample. We proceed to compare our measure to the Census measure of IT jobs by examining state panel regressions in first-differences for 1980, 1990, 2000, 2005, and 2015 where we regress earnings and hours worked growth on growth in the IT jobs' share, conditional on controls. Our objective is to provide additional context for our results.

Table 7 documents our findings. Using our IT intensity measure we estimate that a one percentage point (pp) increase in the share of IT workers is associated with a 0.16pp rise in the growth rate of earnings, whereas a one pp rise in the share of Beckhusen (2016) IT jobs is associated with a 0.06pp rise. Once we add demographic controls we find that the gradient for our measure rises sharply to 0.72pp, whereas the gradient rises for Beckhusen (2016) IT jobs rises only marginally to 0.10pp. We also find that a one pp rise in the our IT intensity share is associated with a 0.12pp rise in the growth rate of hours worked, whereas we find that a one pp rise in the Beckhusen (2016) share is associated with a 0.02pp rise, conditional on controls. Regardless of the IT proxy we use, increases in the share of IT jobs are associated to increases in earnings and hours worked.

Table 7: Comparison of Baseline and Beckhusen (2016) Measures

Dep. var. =	annual earnings growth				annual hours growth			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IT share, growth	.16**		.72***		.21***		.12***	
	[.07]		[.11]		[.02]		[.02]	
Beckhusen share, growth		.06***		.10***		.03***		.02***
		[.01]		[.02]		[.00]		[.00]
R-squared	.02	.10	.48	.49	.38	.31	.67	.71
Sample Size	255	255	255	255	204	204	204	204
Controls	No	No	Yes	Yes	No	No	Yes	Yes

Notes.–Sources: Census Bureau, O*NET, 1980-2015. The table reports the coefficients associated with regressions of state × decade annual earnings and hours worked growth on the share of individuals who are classified as high IT using our baseline O*NET measure and an alternative classification of IT jobs from Beckhusen (2016). Controls (in the labeled specifications) include state growth rates in number of children, family size, age, gender, race, and years of schooling. Standard errors are clustered at the state-level and observations are weighted by the number of observations at a state × period cell.

A.4 A Model of Production with Heterogeneous Tasks

To explicitly account for selection-effects due to unobserved heterogeneity among workers, we develop a model based on the original contribution of Adao (2016). The model formalizes an assignment problem of tasks and technologies in the tradition of Autor (2013).[26] The model features both worker heterogeneity and sorting across tasks, which are identified using the employment shares in different demographic groups, regions, and time periods. Estimates based on this model deliver the implied (shadow) growth rates of IT task prices, relative to non-IT-intensive tasks, as well as estimates of the changes in labor quality composition.

The endogenous selection into sectors and task groups overcomes the conventional concerns associated with the approach of Juhn et al. (1993) since residual wage variation might simply

[26] We adapt Adao's original structure to our setting, deferring readers to Adao (2016) for derivations and a more comprehensive discussion of analytical results.

reflect measurement error or unpriced amenities (Lemieux, 2006). After estimating the model, we (i) examine the distribution of the implied price and worker quality composition effects, (ii) use them to recover, in a second step, the elasticity of substitution between IT and non-IT labor inputs, which we compare to a baseline approach of adjusting for labor services quality.

A.4.1 Task Production

Following Adao's approach, we partition jobs into groups based on geography (state) and college attainment, denoted as $g \in \Im = \{1, 2, .., G\}$, and we allow for mobility between the manufacturing and services sector. We assume that individuals enter the labor market as either college or non-college educated workers.[27] We view states as small open economies with segmented labor markets.[28] Each industry (manufacturing or services) produces its final output by combining two inputs: an IT-intensive and a non-IT-intensive labor aggregate. We assume that these intermediate labor inputs are produced by aggregating individual tasks, denoted j, supplied by different occupations. For simplicity, in the following model description we abstract from industry and we focus on tasks; however, the choice of industry is explicitly accounted for in our empirical analysis.

Workers are heterogeneous in their idiosyncratic productivity at performing different tasks. That is, each worker is endowed with a productivity vector $[L_g^{IT}(i), L_g^{NIT}(i)]$ describing her ability to perform IT and non-IT-intensive tasks.[29] To accommodate the variety of occupations that workers perform, we assume that each aggregate occupation group $k \in \{IT, NIT\}$ is a collection of multiple (perfectly competitive) "single-task occupations" $j \in \Im^k$. Each single task output is used to produce the intermediate labor output k.

The price of the task output of occupation j is denoted as r^j. Like Adao (2016), we assume that workers have identical k-specific productivity within an occupation. That is, workers perform with identical efficiency across all occupations within k. Hence, an occupation j produces task q^j as follows:

$$q^j = Q^j(L_1^j, L_2^j, ..., L_G^j),$$

[27]We do not model the option of returning to school for continuing education.

[28]One could work at a commuting zone level, but we choose not to for two reasons. First, at a lower level of aggregation, our 5% IPUMS sample begins running low on number of observations, producing imprecise shares of IT employment. Second, in talking with geographers and the Census, the classification of commuting zones before and after 2000 has changed quite considerably. While we could work at a county level, which has remained consistently defined over time, our sample size issue would be amplified even further.

[29]In our empirical implementation we allow the productivity vector to vary both by industry and task group. However, in what follows we abstract from industry-specific comparative advantage for ease of exposition.

where $L_g^j = \int_{S_g^j} L_g^{IT}(i)di$ if workers are employed in an IT-intensive occupation and $L_g^j = \int_{S_g^j} L_g^{NIT}(i)di$ if workers are employed in a non IT-intensive occupation. Like Adao, we also assume that the function $Q^j(\cdot)$ is strictly increasing, concave, differentiable and homogeneous of degree one. The technology $Q^j(\cdot)$ allows for the possibility that the labor inputs of different worker groups are imperfect substitutes in production, although this is not required. The integration set S_g^j is the set of workers of type g who are employed in occupation j.

Individual efficiency units can then be used to define a notion of comparative advantage in different occupation groups, i.e. comparative advantage for an individual in the IT sector is $s_g(i) = \ln[L_g^{IT}(i)/L_g^{NIT}(i)]$, whereas absolute advantage is $a_g(i) = \ln[L_g^{NIT}(i)]$. We take individual productivity as exogenous—workers draw their efficiency from a bivariate distribution, i.e. $s_g(i) \sim F_g(s)$ and $\{a_g(i)|s_g(i) = s\} \sim \widetilde{F}_g(a|s)$.

Conditioning on product prices, the labor demand in occupation j of sector k (within group g) is given by:

$$\omega_g^k = r^j \frac{\partial Q^j}{\partial L_g^j} \quad \text{if } j \in \Im^k \tag{8}$$

where ω_g^k is the marginal product of workers in occupation class k and observable group g. Below we estimate how the relative productivity of IT and non-IT-intensive occupations has changed over time, and what these changes suggest about the substitutability of different jobs in the manufacturing and services industries.[30] As in Roy (1951), individuals choose the job that yields the highest utility, which in our case is merely a function of labor income. Letting $y_g^k(i)$ denote the potential logged hourly wage an individual i could earn in sector k, we write:

$$y_g^{NIT}(i) = ln(\omega_g^{NIT}) + a_g(i), \quad y_g^{IT}(i) = \omega_g^{IT} + s_g(i) + a_g(i)$$

where earnings in the IT sector are a function of both comparative and absolute advantage. Because individuals receive different wages based on their comparative advantage, they self-select into the jobs that offer a higher income. Hence, the set of individuals employed in a given sector k can be characterized as:

[30] One potential driver of job substitution is offshoring (Schott, 2004). In fact, any force that shifts the demand for, or supply of, particular jobs is captured in the shadow price changes that we estimate. For example, if non-IT jobs are more likely to be outsourced, one might observe a rising share of IT-intensive jobs and a parallel slowdown in the growth of the shadow price of non IT-intensive jobs.

$$S_g^k \equiv \left\{ i \in \mathcal{I}_g : \ k = \arg\max\{y_g^{IT}(i), y_g^{NIT}(i)\} \right\}$$

Markets are perfectly competitive and wages are such that the demand for labor equals the supply. Since the core part of the model relies on differences in comparative advantage based on individual productivity, we can rank individuals within each group g by their comparative advantage quantile $q \in [0,1]$, so that $\sigma_g(q) \equiv (F_g)^{-1}(q)$ denotes an individual's efficiency in the IT sector based on their rank in the comparative advantage distribution.We can also denote the conditional distribution of absolute advantage as $\widetilde{F}_g(a|\sigma_g(q))$, with an average $\alpha_g(q)$ and variance $v_g(q)$. It follows that the logged wage schedule along the quantile range is :

$$\overline{Y}_g^{NIT}(q) = \ln(\omega_g^{NIT}) + \alpha_g(q), \quad \overline{Y}_g^{IT}(q) = ln(\omega_g^{IT}) + \sigma_g(q) + \alpha_g(q)$$

As shown in Adao (2016), individuals sort into the IT sector if $\sigma_g(q) > \ln(\omega_g^{NIT}) - \ln(\omega_g^{IT})$, otherwise they will sort into the non IT sector. Employment composition is pinned down by marginal individuals with a comparative advantage equal to the relative efficiency-adjusted wage, i.e. $\ln(\omega_g^{NIT}) - \ln(\omega_g^{IT})$, such that $\ln(\omega_g^{NIT}) - \ln(\omega_g^{IT}) = \sigma_g(l_g^{NIT})$ where l_g^{NIT} denotes the employment share in the non IT sector. When we aggregate across the quantiles allocated to a sector, within each group g, we obtain the average logged wage in that sector:

$$\overline{Y}_g^k = \ln(\omega_g^k) + \overline{\alpha}_g^k(l_g^{NIT}) \text{ where } \overline{\alpha}_g^k(l) \equiv \begin{cases} l^{-1} \int_0^l \alpha_g(q)dq & \text{if } k = NIT \\ (1-l)^{-1} \int_l^1 (\sigma_g(q) + \alpha_g(q))dq & \text{if } k = IT \end{cases}$$

A.4.2 Estimating the Growth in Relative Task Prices: IT vs non-IT

For exposition, we abstract from the choice of industry and focus on a bivariate occupation choice (IT vs non-IT). This allows us to illustrate more clearly the way we estimate growth rates in relative returns to different tasks. In the subsequent section, we generalize the approach to account for both occupation and industry choice, so that in the empirical implementation we estimate changes in relative task returns across industries.

We index geographic region by r and let g denote an education-industry group. To obtain an empirical counterpart of the theoretical wage relationships based on comparative advantage, we use the first-order approximation from Adao (2016), which applies to quantiles of the observed wage distribution within each set (g, r, t) of observables. Letting $Y_{g,r,t}(\pi)$ denote the π-th quantile of the log-wage distribution of group g in region r in period t, we estimate

$$\Delta Y_{g,r,t}(\pi) = \Delta\ln(\omega^{IT}_{g,r,t}) + \left[\Delta\ln(\omega^{NIT}_{g,r,t}) - \Delta\ln(\omega^{IT}_{g,r,t})\right] l^{NIT}_{g,r,t_0}(\pi) + \mu_{g,r,t}X_{g,r,t}(\pi) + \Delta\nu_{g,r,t}(\pi) \quad (9)$$

separately for each group-by-region-by-year, where X denotes our usual set of demographic controls. As discussed in Adao (2016), $\Delta\nu_{g,r,t}(\pi)$ is a shock to the absolute advantage of workers in quantile π of the log-wage distribution. Equation 9 identifies $\Delta\ln(\omega^{NIT}_{g,r,t})$ and $\Delta\ln(\omega^{IT}_{g,r,t})$ using the initial sectoral compositions, $l^{NIT}_{g,r,t_0}(\pi)$, and wage growths, $\Delta Y_{g,r,t}(\pi)$, across the quantiles π of the logged wage distribution. This approach leverages variation in the IT employment share across different groups at the start of the sample to identify the task prices growth, $\Delta\ln(\omega^{IT}_{g,r,t})$ and $\Delta\ln(\omega^{NIT}_{g,r,t})$. The intuition behind Equation 9 is that, holding constant the distribution of employees at its initial value $l^{NIT}_{g,r,t_0}(\pi)$, any wage changes that accrue to workers in a given (r, g), between period $t-1$ and t, group must be attributed to changes in relative returns per efficiency unit. The estimator effectively captures that part of the price variation that cannot be attributed to worker composition effects over time.

The identifying assumption in the Adao (2016) approach is that, conditional on $X_{g,r,t}(\pi)$, pre-shock variation in sectoral employment composition is uncorrelated with variation in labor efficiency shocks among individuals with different levels of labor income within the same group-by-region-by-period. The condition is satisfied under the assumptions of a Roy model since marginal workers are indifferent between the two sectors and have similar returns in both. In other words, the assumption more generally implies that reallocation of marginal workers cannot have first-order equilibrium effects on the group's overall wage distribution. A related assumption we make is that there is sectoral reallocation between IT and non-IT jobs and between manufacturing and services jobs; see Adao (2016) for an original discussion of these assumptions.[31]

Having estimated $\Delta\ln(\omega^{IT}_{g,r,t})$ and $\Delta\ln(\omega^{NIT}_{g,r,t})$, one can recover the relative changes in workers'

[31] One concern is that a non-IT-intensive worker in the manufacturing sector may be unlikely to switch into the services sector or re-train with new skills. We mitigate this concern by estimating our model separately for non-college and college workers, as well as controlling for a vector of demographics. More importantly, we examined this concern by computing the gradient between the employment share in manufacturing and the employment share in IT jobs at a county-level separately by five age brackets. We found that the gradients were between -0.40 and -0.46 for workers between ages 25 and 49, but -0.63 for workers above the age of 50. These results somewhat mimic those by Autor et al. (2015), who find that the trade-employment elasticity in manufacturing is coming almost entirely from older workers.

composition by taking the difference between raw wages growth and estimated task prices growth.[32] For each individual (g, r, t) triplet, we are able to recover the price growth and the change in average labor efficiency. These estimates are statistically significant for many of the triplets, although not for all. However, the real value of this procedure lies in the fact that it effectively allows us to approximate the entire distribution of price growth rates for any given group and point in time. Each individual triplet's growth rate can be considered as a (possibly noisy) estimate of a point in the distribution of growth rates for a given industry and group. For our subsequent analysis, we regard the (1980 weighted) mean and median of any such distribution as a reasonable approximation for the growth rate of prices in that industry and group.[33]

A.4.3 Accounting for Selection into Industry

In our empirical analysis, we explicitly account for the possibility that returns to tasks may vary by industry, and that workers may have a different comparative advantage across industries. To this purpose we extend the simple structure described in the previous section to let workers choose between manufacturing and services. In the simple bivariate choice problem discussed in the previous section, the marginal product growth for any given (g, r, t) triplet is a weighted average of the marginal products of workers in IT and non-IT-intensive occupations:[34]

$$\Delta Y_{g,r,t}(\pi) = \Delta \ln(\omega^{IT}_{g,r,t}) l^{IT}_{g,r,t_0}(\pi) + \Delta \ln(\omega^{NIT}_{g,r,t}) l^{NIT}_{g,r,t_0}(\pi) + \Delta \nu_{g,r,t}(\pi). \tag{10}$$

The equation above can be adapted to accommodate more than two types of jobs. The intuition is simple: growth in the marginal product within each job can be recovered by estimating the empirical counterpart of Equation 10 where we consider each job against a combination of all other jobs. We specifically consider four possible jobs for workers to choose from: IT manufacturing, IT services, non-IT manufacturing and non-IT services. In the four jobs case, we can write a set of four equations indexed by industry $i \in \{man, serv\}$ and occupation $k \in \{IT, NIT\}$, as follows:

$$\Delta Y_{g,r,t}(\pi) = \Delta \ln(\omega^{k,i}_{g,r,t}) l^{k,i}_{g,r,t_0}(\pi) + \Delta \overline{\omega}^{k,i}_{g,r,t}(1 - l^{k,i}_{g,r,t_0}(\pi)) + \Delta \nu_{g,r,t}(\pi) \tag{11}$$

[32]Specifically, let $\Delta\kappa^{IT}_{g,r,t}$ and $\Delta\kappa^{NIT}_{g,r,t}$ denote the changes in labor efficiency (composition effects) for IT and non-IT. We can quantify them by solving: $\Delta\kappa^{IT}_{g,r,t} = \Delta \ln(w^{IT}_{g,r,t})/\Delta \ln(\omega^{IT}_{g,r,t})$. To account for outliers and noise, we trim the distributions of both price and composition growth rates at the top and bottom percentiles.

[33]While we have experimented using other weights, we use the number of observations in each cell from 1980. If we, for example, allowed the weights to vary over time, we might confound and/or amplify the role of composition effects.

[34]Adao (2016) provides a derivation of the result from Equation 34 in his job market paper draft.

where $\overline{\omega}^{k,i}_{g,r,t}$ is a weighted average of the growths in logged marginal product of the other three jobs (that is, those that are not in occupation k and industry i). After rearranging, we obtain the following generalized estimation equation, characterizing the marginal product growth of the four job groups available:

$$\Delta Y_{g,r,t}(\pi) = \Delta\overline{\omega}^{k,i}_{g,r,t} + (\Delta\ln(\omega^{k,i}_{g,r,t}) - \Delta\overline{\omega}^{k,i}_{g,r,t})l^{k,i}_{g,r,t_0}(\pi) + \Delta\nu_{g,r,t}(\pi) \tag{12}$$

Since there are two occupation groups k and two industries i, also in this case we have four estimation equations; our next section reports each equation that we use. This specification allows us to recover the growth rate of IT, relative to non-IT, tasks in each of the two industries, for all the triplets (g, r, t).

A.4.4 Generalizing the Task Prices: Industry and Task Groups

In order to extend the model to allow workers to also choose between industries, we begin with the equation 34 of Adao (2016), which enables us to recover the marginal product growth of two types of workers (IT and non-IT in our application):

$$\Delta Y_{g,r,t}(\pi) = \Delta\ln\omega^{IT}_{g,r,t}l^{IT}_{g,r,t_0}(\pi) + \Delta\ln\omega^{NIT}_{g,r,t}l^{NIT}_{g,r,t_0}(\pi) + \Delta\nu_{g,r,t}(\pi) \tag{13}$$

Now, notice that this equation can be reinterpreted to accommodate more than two types of workers. The intuition is the following: the marginal product growth of each kind of job can be recovered by estimating the equation above considering each job against a combination of the other jobs. In the four types of workers case, we can write the following four equations:

$$\Delta Y_{g,r,t}(\pi) = \Delta\ln\omega^{IT,man}_{g,r,t}l^{IT,man}_{g,r,t_0}(\pi) + \Delta\overline{\omega}^{IT,man}_{g,r,t}(1 - l^{IT,man}_{g,r,t_0}(\pi)) + \Delta\nu_{g,r,t}(\pi) \tag{14}$$

$$\Delta Y_{g,r,t}(\pi) = \Delta\ln\omega^{IT,serv}_{g,r,t}l^{IT,serv}_{g,r,t_0}(\pi) + \Delta\overline{\omega}^{IT,serv}_{g,r,t}(1 - l^{IT,serv}_{g,r,t_0}(\pi)) + \Delta\nu_{g,r,t}(\pi) \tag{15}$$

$$\Delta Y_{g,r,t}(\pi) = \Delta\ln\omega^{NIT,man}_{g,r,t}l^{NIT,man}_{g,r,t_0}(\pi) + \Delta\overline{\omega}^{NIT,man}_{g,r,t}(1 - l^{NIT,man}_{g,r,t_0}(\pi)) + \Delta\nu_{g,r,t}(\pi) \tag{16}$$

$$\Delta Y_{g,r,t}(\pi) = \Delta\ln\omega^{NIT,serv}_{g,r,t}l^{NIT,serv}_{g,r,t_0}(\pi) + \Delta\overline{\omega}^{NIT,serv}_{g,r,t}(1 - l^{NIT,serv}_{g,r,t_0}(\pi)) + \Delta\nu_{g,r,t}(\pi) \tag{17}$$

where $\overline{\omega}^{IT,man}_{g,r,t}$, for instance, is a weighted average of the marginal product growth of the three sectors that are not IT manufacturing. Rearranging the four equations above we get the following equations that enable us to estimate the marginal product growth of the four types of jobs that the workers are now allowed to choose:

$$\Delta Y_{g,r,t}(\pi) = \Delta\overline{\omega}_{g,r,t}^{IT,man} + (\Delta \ln \omega_{g,r,t}^{IT,man} - \Delta\overline{\omega}_{g,r,t}^{IT,man}) l_{g,r,t_0}^{IT,man}(\pi) + \Delta\nu_{g,r,t}(\pi) \tag{18}$$

$$\Delta Y_{g,r,t}(\pi) = \Delta\overline{\omega}_{g,r,t}^{IT,serv} + (\Delta \ln \omega_{g,r,t}^{IT,serv} - \Delta\overline{\omega}_{g,r,t}^{IT,serv}) l_{g,r,t_0}^{IT,serv}(\pi) + \Delta\nu_{g,r,t}(\pi) \tag{19}$$

$$\Delta Y_{g,r,t}(\pi) = \Delta\overline{\omega}_{g,r,t}^{NIT,man} + (\Delta \ln \omega_{g,r,t}^{NIT,man} - \Delta\overline{\omega}_{g,r,t}^{NIT,man}) l_{g,r,t_0}^{NIT,man}(\pi) + \Delta\nu_{g,r,t}(\pi) \tag{20}$$

$$\Delta Y_{g,r,t}(\pi) = \Delta\overline{\omega}_{g,r,t}^{NIT,serv} + (\Delta \ln \omega_{g,r,t}^{NIT,serv} - \Delta\overline{\omega}_{g,r,t}^{NIT,serv}) l_{g,r,t_0}^{NIT,serv}(\pi) + \Delta\nu_{g,r,t}(\pi) \tag{21}$$

A.5 Elasticity Estimates: Alternative IT Intensity Measure

In the paper we report estimates of the elasticities based on the baseline definition of IT intensity, which relies on the average z-score of IT intensity. An alternative modeling choice would be to define IT intensive occupations based on whether their IT score exceeds the median IT intensity. For ease of comparison Table 8 reports elasticity estimates under both definitions of high IT intensity: the baseline (mean) and the median.

Estimates are not substantially different and suggest that the value we estimate are robust to reclassifying occupations in terms of their IT intensity.

	Quality-Adjusted Measures						Quality & Selection -Adjusted Measures			
	(M)	(S)	(M)	(S)	(M)	(S)	(M)	(S)	(M)	(S)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Baseline	1.72***	1.24***	1.28***	1.15***	1.31***	1.38***	1.63***	1.31***	1.77***	1.56***
	(0.138)	(0.049)	(0.066)	(0.045)	(0.115)	(0.116)	(0.056)	(0.042)	(0.092)	(0.086)
Median	1.38***	1.55***	1.24***	1.22***	1.27***	1.36***	1.68***	1.51***	1.72***	1.85***
	(0.083)	(0.077)	(0.060)	(0.051)	(0.097)	(0.112)	(0.051)	(0.058)	(0.074)	(0.118)
Time Diff.(Δ)	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument	No	No	No	No	Yes	Yes	No	No	Yes	Yes
N	408	408	306	306	306	306	267	267	267	267

Table 8: Comparison of Elasticity Estimates: Alternative Measure of IT Intensity

Notes.–Sources: Census Bureau, 1990-2015. The table reports the elasticity of substitution between information technology (IT) and non-IT workers in two ways. Columns 1-6 estimate the elasticity according to the following equation

$$\Delta \ln \left(\omega_i^{IT} L_i^{IT} / \omega_i^{NIT} L_i^{NIT}\right) = \Delta_t \ln \left(\alpha_i/(1-\alpha_i)\right) + \nu_i \Delta_t \ln \left(L_i^{IT}/L_i^{NIT}\right) + \nu_i \Delta_t \left(\varepsilon_i^{IT}/\varepsilon_i^{NIT}\right)$$

where our outcome variable is the ratio of the wage bills and our right-hand-side variable is the ratio of effective labor services, which we estimate in log-levels and first-differences, using ordinary least squares and instrumental variables. We obtain quality-adjusted labor services premia by first estimating a group $\times$ year skill price obtained by regressing annual earnings on an indicator for working in an IT job, conditional on controls, and subsequently dividing total earnings by the skill price proxy and multiplying by an individual's total hours worked during the year. Controls in the skill price estimation include race, family size, marital status, gender, a quadratic in years of schooling, and age bin fixed effects (20-29, 30-39, 40-49, 50+). Controls in our estimation in the equilibrium condition include: group-specific average family size, share of males, race shares, a quadratic in schooling, and age bin shares. Columns 7-10 estimate the elasticity according to the following equation

$$\Delta \ln \left(\omega_i^{IT} / \omega_i^{NIT}\right) = \Delta_t \ln \left(\alpha_i/(1-\alpha_i)\right) + (\nu_i - 1) \Delta_t \ln \left(L_i^{IT}/L_i^{NIT}\right) + \nu_i \Delta_t \left(\varepsilon_i^{IT}/\varepsilon_i^{NIT}\right)$$

where the difference from columns 1-6 is that our skill prices and quality-adjusted measure of labor services are obtained from our selection model based on Adao (2016). Controls in all specifications are the same and our instruments include the IT employment premia growth between 1970 and 1980 and the share of IT workers in 1970. Observations are weighted by the number of workers observed in each cell in 1980.The only difference between the "mean" and "median" rows are that the "median" elasticities are from measures of high IT defined by whether the individual is above the median IT intensity z-score, whereas the "mean" from our baseline is defined by whether the individual is above the mean (z-score of zero by construction).