Heterogeneous Effects of Capital-Embodied Innovation on Labor Market *

Hyejin Park† Younghun Shim‡

June 2025

Abstract

This paper develops an occupation-level measure of Capital-Embodied Innovation (CEI) by matching patents with capital goods based on their text similarity. The impact of CEI on labor demand is heterogeneous, depending on the similarity between capital and occupational tasks. Specifically, CEI associated with task-similar capital reduces the relative labor demand, whereas CEI related to task-dissimilar capital raises it. Between 1980 and 2015, abstract and non-routine occupations experienced more innovations in task-dissimilar capital and fewer in task-similar capital. CEI can explain a significant fraction of the changes in the task-biased labor market and the decrease in labor share.

Keywords: Capital-Embodied Innovation, Text Analysis of Patents, Substitution between Labor and Capital
JEL codes: J24, J31, O33, O47

*We thank Ufuk Akcigit, Julieta Caunedo (discussant), Erik Hurst, Greg Kaplan, Keun Lee, Josh Morris-Levenson, Simon Mongey, Robert Shimer, Yongseok Shin, Nancy Stokey, and seminar participants at Southern Economic Association 2023, Korea Development Institute 2023 Workshop, NBER Summer Institute Meeting 2023, NASMES 2023, CESifo 2023, Midwest Macro 2023, Chung-Ang U., Korea U., SNU, Yonsei, SED 2022, Hong Kong Baptist U. for their helpful comments. The paper has been previously circulated under the title "How Task-Biased is Capital-Embodied Innovation?" The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management. The authors gratefully acknowledge the financial support from the Social Sciences and Humanities Research Council in Canada [PV153480, PVX20020]. All errors are ours.

†Université de Montréal; Email Address: hyejin.park@umontreal.ca
‡International Monetary Fund; Email Address: yshim@imf.org

1 Introduction

New technologies have raised concerns about unequal labor market outcomes for centuries. For instance, the invention of cotton-spinning machinery in the 19th century displaced handicraftsmen while increasing demand for machine operators. Recent technologies, such as computers and artificial intelligence, are under intense scrutiny for their potential to displace many occupations while benefiting others. Accurate and systematic measurement of these innovations is essential to understand their impacts and inform public policy.

Previous papers addressing this issue have focused on a few significant episodes of new technologies, often embedded in capital, and measured their varied exposures. Autor and Dorn (2013) analyze the rise of computers that substitute routine tasks, while Acemoglu and Restrepo (2022) investigate the role of robots that replace workers in manufacturing industries. However, focusing on a few types of capital goods can overlook a significant fraction of capital that reflects innovation. Robots and computers accounted for a small fraction of capital expenditure, with 0.7% and 3% of equipment expenditures in 2019, respectively.[1] Therefore, a wider range of capital needs to be examined to capture innovation embodied in capital more accurately.

This paper constructs a measure of capital-embodied innovation (CEI) across a comprehensive set of capital goods at the occupation level and examines its heterogeneous effects across different occupations. We first use O*NET to assign capital goods to occupations and classify them into two types based on their similarity to the occupational tasks. If a capital good performs a function similar to the tasks of an occupation, it is classified as *task-similar* for that occupation. Conversely, if the function of a capital good differs from occupational tasks but is still used by occupations, it is classified as *task-dissimilar*. This classification is made by computing text similarity scores between descriptions of capital goods from Wikipedia and occupational task statements from O*NET. Then, CEI is measured by matching patents with capital goods based on text similarities between

[1] The computer expenditure share is from BEA fixed assets, and the robot share is from the 2019 Annual Capital Expenditure Survey of the Census Bureau. Even when combined with related equipment, such as mainframes and storage devices, computer-related equipment makes up 9.7% of total capital expenditures.

abstracts of patents and Wikipedia articles on capital goods. Equipped with this measure, we quantify the role of CEI in changes in the labor market across occupations using a structural model.

Linking new technologies with capital offers a bridge between the literature on capital-embodied technical changes and the labor-market effects of innovation. Caunedo et al. (2023) recently evaluate the technical changes embodied in capital at the occupation level based on capital prices. Our method helps identify a source of declines in capital prices associated with new technologies. Changes in capital prices can be due to various factors, such as innovation, trade, and changes in market structure. Isolating technological factors behind these price changes is crucial for recent discussions on the impact of R&D subsidies on inequality (Bloom et al., 2019).

At the same time, by matching new technologies with capital, we can distinguish whether new technologies affect labor demand by accelerating substitution toward capital or increasing demand for occupation services. Kogan et al. (2023) and Autor et al. (2024) focus on patents closely related to occupational tasks or micro-titles. Our analysis extends this approach by introducing capital goods as intermediaries between innovation and labor. By doing so, we can measure how the price and quantity of capital goods change for each occupation along with the wage or employment. Whether innovation promotes substitution with capital or the overall production of occupational services yields different implications on the labor share.

Furthermore, using capital goods as intermediaries allows us to link more specific technologies to each occupation. Rather than matching patents directly to occupational tasks or micro-titles, as in Kogan et al. (2023) and Autor et al. (2024), we leverage the occupation-level list of capital goods to identify which technologies are relevant for each occupation. Occupational task descriptions are often broad, using terms like "controlling machines and processes", which can obscure important differences. For example, interior designers and fashion designers have identical task descriptions, yet their capital goods differ substantially: only 6 out of 16 capital goods used by interior designers are also used by fashion designers. For example, using "sewing machines" as intermediaries, patents related to sewing technologies can be accurately identified as relevant innovations for fashion designers.

We begin by building a general equilibrium model in which occupational service is produced using task-similar and task-dissimilar capital alongside occupational labor. The two types of capital are allowed to have different elasticities of substitution with labor. Depending on the relative magnitude of the elasticity of substitution, changes in the cost of capital can increase or decrease the demand for labor at the occupation level.

The parameters of this model are estimated using a linear regression equation derived from the first-order conditions of cost minimization for occupational service production. To identify the effect of CEI alongside elasticities of substitution, we devise shift-share instruments from academic publications, capital imports, and immigration trends from Latin American countries. An increase in imported capital goods raises the capital expenditure share in occupations that intensively use those goods, while increased immigration from Latin America increases labor supply disproportionately in occupations where the immigrants have a comparative advantage. Additionally, a rise in publications within an academic field spurs innovation activities in related patent classes.

In our framework, CEI affects occupational labor demand in two ways: by reducing the user costs of capital and impacting the demand for occupational services. The estimated elasticities of substitution suggest that lower user costs associated with CEI in task-similar capital reduce occupational labor demand, whereas those in task-dissimilar capital increase it. Furthermore, CEI in task-similar capital decreases demand for occupational services, while CEI in task-dissimilar capital increases it.

We find that between 1980 and 2015, occupations were heterogeneously exposed to CEI. First, the magnitude of CEI varied between occupations. CEI in task-dissimilar capital (CEI-d) was biased toward abstract and non-routine occupations with high labor shares and wages, whereas CEI in task-similar capital (CEI-s) favored non-abstract occupations with low labor shares and wages. Second, occupations experienced changes in capital intensity, altering the impacts of CEI. Non-abstract, routine occupations with low labor shares became relatively more intensive in task-similar capital, while abstract, high-wage occupations became more intensive in task-dissimilar capital.

To quantify the role of CEI on various labor market trends, we conduct counterfactual exercises by fixing the measure of patents at their levels in 1980. First, we examine the role of CEI in task-biased changes. Our results suggest that CEI accounts for 18–59% of employment growth and almost the entire wage growth that favors abstract occupations. Likewise, CEI produces 8–27% and 70–79% of the bias against routine occupations in wage and employment growth, respectively. Second, CEI contributes to the decline in labor share, generating 86–89% of the observed decline between 1980 and 2015. Lastly, CEI helps explain job polarization, accounting for 7–27% of employment growth and 72–79% of wage growth for high-wage occupations in the top quintile of the 1980 wage distribution.

Related Literature

This paper first contributes to the literature on task-biased technical changes and job polarization (e.g., Autor et al., 2006; Goos and Manning, 2007; Lee and Shin, 2017; Bárany and Siegel, 2018; Keller and Utar, 2023). Most papers in this literature, including Autor and Dorn (2013), Goos et al. (2014), Michaels et al. (2014), and Acemoglu and Restrepo (2022), study the emergence of specific capital goods such as computers, information technology equipment, and robots. They find that new technologies in these capital goods have reduced the demand for routine and non-abstract occupations. Since middle-wage occupations are more likely to be routine, these changes contributed to job polarization. Our research extends these works by developing a measure of innovation for a comprehensive range of capital goods at the occupational level. Additionally, this paper distinguishes technological factors from other drivers of capital goods prices, which helps to understand the uneven impacts of innovation policies.

Second, this paper relates to the broader literature that studies the complementarity between capital and worker skills (Griliches, 1969; Goldin and Katz, 2008; Hornstein et al., 2005). Most papers assume that workers from different skill groups have different elasticities of substitution with capital, and the magnitude of elasticity determines how labor demand for a worker group responds to capital accumulation (Krusell et al., 2000; Berlingieri et al., 2022; Caunedo et al., 2023). In contrast, our analysis categorizes capital goods into two types and allows these types to have different elasticities of substitution with labor. This approach captures a rich heterogeneity in complementarity between capital and labor with only two elasticities of substitution.

Lastly, this paper contributes to a growing literature that applies textual analysis to patent data to measure innovation (Argente et al., 2023; Hémous et al., 2025; Zhestkova, 2021; Bloom et al., 2021; Kelly et al., 2021; Mann and Püttmann, 2023). Many existing papers match patents similar to task descriptions of occupations to measure exposure to new technologies. Webb (2019) matches occupations with technologies related to artificial intelligence and robots, while Kogan et al. (2023) include a broader set of new technologies for matching. Autor et al. (2024) categorize labor-augmenting and labor-saving technologies by matching patents with micro titles and tasks of occupations. Recently, Hémous et al. (2025) use patent texts to identify patent classes related to automation technologies. We introduce "Tools Used" data from O*NET to match patents with capital goods used by occupations. This approach allows our innovation measure to include new technologies not similar to occupational tasks but utilized by occupational workers in the form of capital. Our results indicate that these technologies also reallocate labor demand between occupations and are quantitatively as important as new technologies similar to occupational tasks.

The remainder of the paper is organized as follows. Section 2 outlines the empirical framework. Section 3 describes the data used for the analysis and the procedure to construct CEI measures. Section 4 discusses the estimation strategy and the results. Section 5 presents the results from counterfactual exercises. Section 6 concludes.

2 Empirical Framework

2.1 Overview

The economy is static and consists of firms and workers. Final goods are produced using industrial outputs. A representative firm in each industry combines occupational services to create industrial outputs. For example, an aerospace company integrates occupational services from aerospace engineers, engine mechanics, and janitors to produce its industrial output. These occupational services are produced with labor and capital, where capital is a bundle of individual capital goods. The labor of engine mechanics, for example, is combined with capital bundles that include tools such as pressure indicators and wire cutters.

Two types of capital enter the production of occupational services. First, task-similar capital consists of tools that perform similar functions as occupational tasks. In contrast, capital goods in task-dissimilar capital bundles fulfill functions distinct from occupational tasks but essential to producing occupational services. One capital good can be task-similar for one occupation but task-dissimilar for another. For instance, an engine test stand is considered task-similar capital for engine mechanics involved in engine maintenance. However, the same engine test stand is categorized into task-dissimilar capital for aerospace engineers designing new aircraft. We allow these two types of capital to have different elasticities of substitution with labor.

Capital is bundled from individual capital goods and is supplied elastically at the unit cost, which is described below. Different occupations work with capital bundles with different compositions of capital goods. In addition, each industry requires a different mix of capital goods for a given occupation. Thus, the composition and user costs of capital bundles vary by occupation and industry.

The labor market is distinguished by occupations but not by industries. Thus, the wage is set at the occupation level, and workers are indifferent across industries within an occupation. Workers select the occupation that offers them the highest utility, considering wages and idiosyncratic preferences. Firms in each industry hire workers of different occupations. The equilibrium occupational wages clear all occupational labor markets.

In this economy, CEI shifts occupational labor demand by changing the user costs of capital goods and, thereby, the user costs of capital bundles. Moreover, CEI directly shifts the relative demand for occupational services. This assumption is motivated by the idea that innovation can alter the role of occupational services in industrial production, regardless of capital costs.

2.2 Capital Bundle

Competitive capital producers combine capital goods to make occupation- and industry-specific bundles of task-similar and task-dissimilar capital, k_{jio}, where $j \in s, d$ denotes capital type, with s representing task-similar capital and d representing task-dissimilar capital, for industry i and occupation o. Capital goods are combined to produce k_{jio} as follows:

$$k_{jio} = f(x_{jio1}, \cdots, x_{jioN}),$$

where x_{jion} is the quantity of capital goods n, and $f(\cdot)$ is a constant-return-to-scale aggregator.[2]

The user cost of the capital bundle is determined by the zero profit condition.

$$r_{jio} = \sum_{n \in \mathbb{N}_{jo}} \lambda^k_{jion} \frac{x_{jion}}{k_{jio}}, \tag{1}$$

where λ^k_{jion} is the user cost of capital goods, and $\mathbb{N}_{jo}$ is a set of capital goods that are categorized as group j for occupation o. Notice that this condition holds whenever the production of a capital bundle exhibits a constant-returns-to-scale property, the zero-profit condition holds, and x_{jion}/k_{jio} is the quantity share of capital goods n.

Later, we will measure CEI using the number of patents matched to a capital good used by an occupation in an industry, # Patent$_{jon}$, which will be described in more de-

[2] In the counterfactual analysis, we assume $f(\cdot)$ is a CES aggregator with the elasticity of substitution with $\phi = 1.13$, following Caunedo et al. (2023).

tail in Section 3.3. If patents change the user costs of capital goods isoelastically, i.e., $d\log\lambda^k_{jion} \propto \gamma_P d\log\left(\#\,\text{Patent}_{jon}\right)$, we can take the total derivative of r_{jio} to get the following equation.

$$d\log r_{jio} = \gamma_P \underbrace{\sum_n \frac{\lambda^k_{jion} x_{jion}}{r_{jio} k_{jio}} d\log\#\,\text{Patent}_{jon}}_{=:d\log P_{jio}} \tag{2}$$

Equation (2) defines the occupation-level CEI measure, $d\log P_{jio}$. We approximate $d\log P_{jio}$ with discrete differences between 1980 and 2015 as in the following equation.

$$d\log P_{jio} \approx \Delta\log P_{jio} \equiv \sum_n \bar{\omega}_{jion}\Delta\log\#\,\text{Patent}_{jon} \tag{3}$$

In this equation, $\bar{\omega}_{jion}$ is the average expenditure share on capital goods between 1980 and 2015, and $\Delta\log\#\,\text{Patent}_{jon}$ is the difference in the number of patents between 1980 and 2015. A negative γ_p implies that the user costs of capital decrease with CEI. It is important to note that the coefficient γ_p does not vary across similar and dissimilar capital groups. From now on, the average patent change for the capital bundle, $\Delta\log P_{jio}$, is the measure of CEI-j, where $j = s$ for similar capital and $j = d$ for dissimilar capital.

2.3 Labor Demand

Aggregate output Y is a Cobb-Douglas composite of industrial outputs.

$$\mathbf{Y} = \prod_i Y_i^{\alpha_i}.$$

Industrial output in industry i, Y_i, aggregates occupational services with a constant elasticity of substitution, σ.

$$Y_i = \left(\sum_o \mu_{io}^{\frac{1}{\sigma}} y_{io}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \tag{4}$$

where μ_{io} is the occupation demand shifter for industry i, occupation o. Occupational service y_{io} is produced with capital and labor as in the following equations.

$$\Theta_{io} = \left(z_{sio}^{\frac{1}{\rho_s}} k_{sio}^{\frac{\rho_s - 1}{\rho_s}} + l_{io}^{\frac{\rho_s - 1}{\rho_s}} \right)^{\frac{\rho_s}{\rho_s - 1}}, \tag{5}$$

$$y_{io} = \left(z_{dio}^{\frac{1}{\rho_d}} k_{dio}^{\frac{\rho_d - 1}{\rho_d}} + \Theta_{io}^{\frac{\rho_d - 1}{\rho_d}} \right)^{\frac{\rho_d}{\rho_d - 1}}. \tag{6}$$

In these equations, k_{sio} denotes task-similar capital with its productivity, z_{sio}, and k_{dio} is task-similar capital with its productivity, z_{dio}. l_{io} denotes labor inputs in industry i and occupation o. As in Krusell et al. (2000), the nested CES structure allows different substitutability between production inputs. ρ_s and ρ_d are the elasticity of substitution of labor with task-similar and task-dissimilar capital, respectively. This structure implies that the elasticity of substitution between task-dissimilar capital and task-similar capital is also ρ_d.[3]

A representative firm of industry i chooses labor and capital inputs to minimize the production costs, given the user costs of task-similar and task-dissimilar capital, r_{sio} and r_{dio}, and the occupational wage w_o. The first-order conditions are described as the following equations.

$$\frac{r_{sio}}{w_o} = z_{sio}^{\frac{1}{\rho_s}} \left(\frac{k_{sio}}{l_{io}} \right)^{-\frac{1}{\rho_s}}, \tag{7}$$

$$\frac{r_{dio}}{w_o} = \Theta_{io}^{\frac{\rho_s - \rho_d}{\rho_s \rho_d}} z_{dio}^{\frac{1}{\rho_d}} k_{dio}^{-\frac{1}{\rho_d}} l_{io}^{\frac{1}{\rho_s}}, \tag{8}$$

$$\frac{w_o}{w_{o'}} = \left(\frac{\mu_{io}}{\mu_{io'}} \right)^{\frac{1}{\sigma}} \left(\frac{y_{io}}{y_{io'}} \right)^{-\frac{1}{\sigma} + \frac{1}{\rho_d}} \left(\frac{\Theta_{io}}{\Theta_{io'}} \right)^{\frac{\rho_d - \rho_s}{\rho_s \rho_d}} \left(\frac{l_{io}}{l_{io'}} \right)^{-\frac{1}{\rho_s}}. \tag{9}$$

Combining these three equations gives the following equation that governs the relative

[3] In Appendix D.2, we explore an alternative specification in which labor and task-dissimilar capital are aggregated first, followed by aggregation with task-similar capital. The estimation results remain quantitatively similar.

labor demand within industry i.

$$\frac{w_o}{w_{o'}} = \left(\frac{\mu_{io}}{\mu_{io'}}\right)^{\frac{1}{\sigma}} \left(\frac{\tilde{y}_{io}}{\tilde{y}_{io'}}\right)^{-\frac{1}{\sigma}+\frac{1}{\rho_d}} \frac{\tilde{\Theta}_{io}^{\frac{\rho_d-\rho_s}{\rho_s\rho_d}}}{\tilde{\Theta}_{io'}^{\frac{\rho_d-\rho_s}{\rho_s\rho_d}}} \left(\frac{l_{io}}{l_{io'}}\right)^{-\frac{1}{\sigma}}. \tag{10}$$

In these equations, $\tilde{\Theta}_{io} = \Theta_{io}/l_{io}$ and $\tilde{y}_{io} = y_{io}/l_{io}$ are defined as the labor efficiencies for the inner and the outer composites of occupational service production. In equilibrium, they can be expressed as follows.

$$\tilde{\Theta}_{io} = \left(z_{sio}\left(\frac{r_{sio}}{w_o}\right)^{1-\rho_s} + 1\right)^{\frac{\rho_s}{\rho_s-1}}, \tag{11}$$

$$\tilde{y}_{io} = \tilde{\Theta}_{io}^{\frac{\rho_s-\rho_d}{\rho_s}} \left(z_{dio}\left(\frac{r_{dio}}{w_o}\right)^{1-\rho_d} + \tilde{\Theta}_{io}^{\frac{\rho_d-1}{\rho_s}}\right)^{\frac{\rho_d}{\rho_d-1}}. \tag{12}$$

$\tilde{\Theta}_{io}$ and $\tilde{y}_{io}$ decrease unambiguously with r_{sio} and r_{dio}, respectively. In other words, lower user costs of capital increase the labor efficiencies for the inner and outer composites of occupational service production.

Equation (10) shows that the relative magnitudes of the elasticities of substitution shape how capital-embodied changes affect labor demand between occupations, consistent with Caunedo et al. (2023). A decrease in user costs of task-dissimilar capital increases $\tilde{y}_{io}$ and raises demand for occupational services through scale effects. If $\sigma > \rho_d$, the demand rises more elastically than the substitution toward task-dissimilar capital, increasing the relative labor demand.

Likewise, if $\rho_s > \sigma$, the substitution toward task-similar capital is stronger than the overall demand increase for occupational services. An increase in $\tilde{\Theta}_{io}$ from lower user costs of task-similar capital raises both $\tilde{y}_{io}$ and $\tilde{\Theta}_{io}$. Since $d\log\tilde{y}_{io}/d\log\tilde{\Theta}_{io} < 1$, $\rho_s > \sigma$ implies that lower user costs of task-similar capital reduce relative labor demand. Thus, in this framework, the effect of CEI on user costs of capital depends on the relative magnitudes of elasticities between capital and labor, as well as across occupational services.

We also allow the CEI to directly affect the demand for occupational services. We assume that the same CEI measure in Equation (2), $d\log P_{jio}$, changes the demand shifter, μ_{io}, as in the following equations.

$$d\log \mu_{io} = \gamma_s d\log P_{sio} + \gamma_d d\log P_{dio}. \tag{13}$$

A positive γ_j implies that the demand for occupational services increases with CEI-j ($j \in \{s,d\}$), even after taking into account its effect through substitution with capital in Θ_{io} and $\tilde{y}_{io}$. This can happen when the quality of occupational service increases or when the scope of the production process implemented by an occupation increases with CEI. Additionally, the demand shifter margin addresses the effect of CEI that the nested CES structure cannot capture.

2.4 Labor Supply and Equilibrium

The labor supply side is modeled with a standard structure of occupation choice. Let $\mathbf{L}$ denote the number of ex-ante homogeneous workers. Workers observe the wage of each occupation determined in the market, w_o, occupation-specific utility ξ_o, and idiosyncratic utility realized for each occupation ζ. The worker chooses an occupation that gives the highest utility. All workers receive the same wage and utility for any given occupation. Consequently, once they choose an occupation, they are indifferent across industries. The occupation choice problem can be written as follows:

$$o^* = \underset{o}{\operatorname{argmax}} \left\{ \log w_o + \log \xi_o + \zeta \right\}.$$

Assuming that ζ follows an i.i.d. Type 1 Extreme Value Distribution with scale parameter $1/\eta$, the following equation determines the supply of occupational labor.

$$\frac{L_o}{\mathbf{L}} = \frac{\exp(\eta \log w_o + \eta \xi_o)}{\sum_{o'} \exp(\eta \log w_{o'} + \eta \xi_{o'})}. \tag{14}$$

The labor market equilibrium consists of occupational wages that equate the supply of labor with the demand in each occupational labor market, which consists of demands at the industry level for each occupation.

3 Data and Measurement

3.1 Data

The data from O*NET "Tools Used" serve as our primary reference for identifying capital goods with which each occupation works.[4] O*NET compiles a comprehensive list of machines or equipment vital to occupational roles (Dierdorff et al., 2006). To illustrate, security managers use capital goods such as security control systems, alarm systems, and video monitors. The data encompass 4,180 distinct capital goods used by 775 occupations coded in the 2010 Standard Occupational Classification Code (SOC). In particular, each capital good is associated with a title and a corresponding United Nations Standard Products and Services Code (UNSPSC).

We use patent data from the United States Patent and Trademark Office (USPTO) to measure innovation on these capital goods.[5] This dataset includes all the patents registered in the US from 1970 to 2015. The exercise uses the application year, title and abstract of patents.[6] In total, we have 6.1 million utility and plant patents. Design patents are excluded to focus on quality improvement.

For occupational employment at the industry level in 1980 and 2015, we used microdata from the 1980 Decennial Census and the American Community Survey (ACS) from 2015 to 2019 for observations in 1980 and 2015, respectively. The data is downloaded from the Integrated Public Use Microdata Series (IPUMS). ACS samples from multiple surveys

[4] This study uses version 25.0, which was updated in August 2020, whereas Caunedo et al. (2023) use the Dictionary of Occupational Titles and different versions of "Tools Used" to capture changes in capital goods at the occupation level. We set the version of the "Tools Used" data to be consistent with the Wikipedia data, which is also from 2020. IV estimation addresses the measurement errors associated with changes in capital goods over time.

[5] The bulk file is downloaded from `https://patentsview.org`.

[6] The application year is used instead of the grant year since it is closer to the actual innovation year.

are used to increase the size of the samples in each occupation by industry. Employment is measured by the number of workers with the occupation and industry codes in the 1990 Census classification system harmonized by the IPUMS. Each observation is used with sampling weights from the Census Bureau. Our analysis uses prime-aged workers between the ages of 25 and 54. The Decennial Census and the ACS are also used to construct immigrant supply instruments in Section 4.

Occupational wages are sourced from the microdata for the Annual Social and Economic Supplement (CPS-ASEC) of the Current Population Survey. The wage is measured by the average weekly labor earnings and computed as the annual labor income divided by the number of weeks worked. Observations in 1980–1984 and 2015–2019 are used to calculate wages in 1980 and 2015, respectively.[7]

To account for heterogeneous labor productivity between workers with different observable characteristics, we residualize wages using Mincerian regression, which includes age, education level, race, and year dummies, as in Berlingieri et al. (2022). For this regression, we only consider full-time male workers who worked 40 weeks or more in the preceding year. Samples with zero or missing information on individual characteristics are excluded. Furthermore, observations with a nominal hourly wage below 50% of the federal minimum wage for the given year are omitted.

The 2010 SOC on the O*NET data is mapped to the OCC1990 variable using correspondence between the OCC1990 variables and the 2010 SOC in the ACS 2012-2018. Likewise, the IND1990 variable is converted to the NAICS code using the correspondence between the IND1990 and the NAICS in the ACS. Then, the NAICS in the ACS is aggregated to the 63 NAICS industries in the National Income and Product Accounts (NIPA) by the Bureau of Economic Analysis (BEA).

For capital stocks and user costs of capital at the occupation and industry level, we use fixed- and current-cost capital estimates from the BEA. Fixed-cost estimates are measured

[7]The CPS-ASEC is not used to measure employment at the occupation and industry level because of its small sample size. The CPS-ASEC is not used to measure employment at the occupation and industry level because of its small sample size. We do not use the wage variables from the ACS and the Decennial Census because the data do not include occupation information for last year.

in the 2012 US dollar at the industry and the NIPA capital category level. Fixed-cost capital serves as the quantity of capital goods, and the ratio between current- and fixed-cost estimates is used to measure the price of capital goods in the calculation of user costs.

We introduce additional data sets to construct instrumental variables with academic publications and capital imports, which are described in more detail in the Appendix C. First, we use patent citations to academic publications from Marx and Fuegi (2020) and the number of publications from Microsoft Academic Graph (MAG, Sinha et al., 2015). Combining these two datasets allows us to gauge the knowledge flow from academic research to patents and offers variations that affect patenting activities. Second, we work with UN Comtrade data to measure import volume at the commodity level.[8] This information helps us capture increases in capital supply at the capital good level from international trade.

3.2 Identifying Task-Dissimilar versus Task-Similar Goods

For each occupation, we categorize the capital goods into two categories: task-similar capital and task-dissimilar capital. The capital whose function closely aligns with the tasks of an occupation is categorized as task-similar. In contrast, the capital used by an occupation whose function does not mirror the occupational tasks is labeled as task-dissimilar. One capital good may be task-similar for one occupation and task-dissimilar for another, reflecting the heterogeneity of tasks across various occupations. At this point, we only allow for different degrees of substitution elasticity between the two types of capital and labor and do not presuppose their relationships with occupational labor demand before the estimation.

The existing literature that matches occupations with patents based on text similarity (e.g., Webb, 2019; Kogan et al., 2023) often finds a strong labor-displacement effect of innovations. Our classification is motivated by the negative effect of new technologies

[8]Comtrade data for 1980 and 2015 are available at the SITC Rev. 2 and HS 1992 levels, respectively. We converted SITC Rev. 2 in HS 1992 using a crosswalk file provided by the UN Statistics Division and then manually converted it to the NIPA code.

that perform tasks similar to those of occupations. If new technologies conduct functions that differ from occupational tasks but are used by the occupation, these technologies may have different effects. The occupation-level list of capital goods provided by O*NET serves as a convenient intermediary to identify the new technologies used by each occupation.

Specifically, the classification exploits the degree of similarity of text between the tasks associated with an occupation and the descriptions of capital goods. We use data from "Task Statements" in O*NET for occupational tasks.[9] For example, a security manager has tasks such as "Respond to medical emergencies, bomb threats, fire alarms, or intrusion alarms, following emergency response procedures." For descriptions of capital, we use Wikipedia articles, which offer product-level descriptions for text analysis (Argente et al., 2023). Utilizing the Wikipedia Application Programming Interface (API), we locate Wikipedia pages for 1,825 among 4,180 capital goods listed.[10]

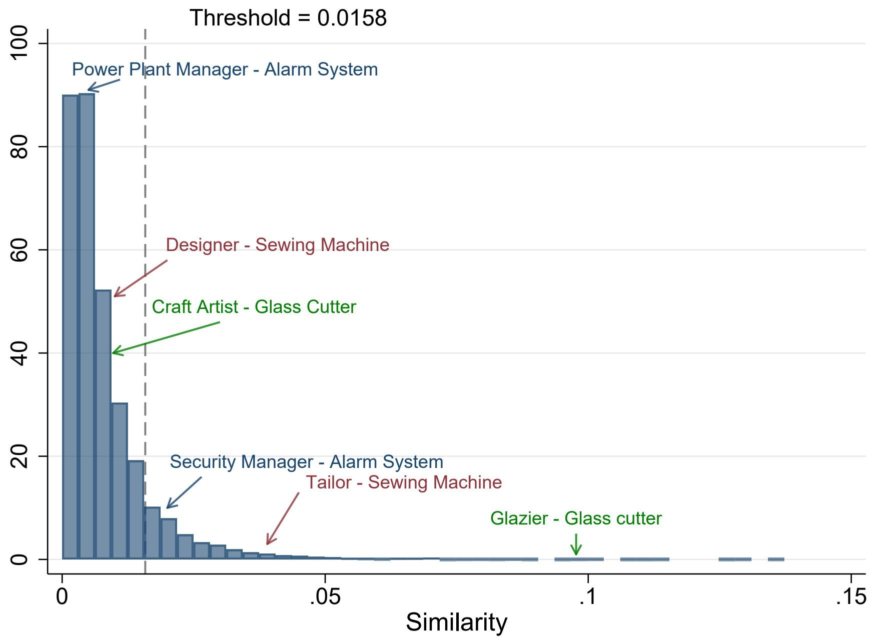
We then compute the text similarity between Wikipedia articles describing capital goods and occupational tasks by counting the common words. A standard procedure from the natural language processing literature is used to prepare the texts for our analysis. First, we remove stopwords, words that are insignificant in delivering the content. For example, "is," "where," and "have" are classified as stopwords. Removing them prevents erroneous matches between two texts solely based on shared functional words rather than substantive content. Then, words are lemmatized to standardize word forms. For example, "generating" or "generated" is changed to "generate." This step ensures that words with analogous meanings, though in different forms, align appropriately.

Next, we calculate the pairwise similarity between tasks and capital goods. Specifically, each text is vectorized to compute cosine similarity, which quantifies the share of overlapped words between two texts. Words are weighted by the frequency-inverse document frequency (TF-IDF). The weight of words $\mathbf{i}$ in document $\mathbf{j}$, represented as $\mathbf{w_{ij}}$, is

[9] We use the 25.0 version, updated in August 2020. On average, each occupation has 23 tasks.

[10] The Wikipedia data was downloaded on 02/28/2021. Table A1 details the proportion of tools found in Wikipedia, categorized by their NIPA category. Tools related to electronics, furniture, and machinery are more frequently found, while those related to mining, medical equipment, and aircraft are less common.

Figure 1: Distribution of Similarity Scores



Notes. This figure plots the density of similarity scores for the pairs of capital goods and occupations. The text similarity score is initially measured at the task level for each capital good and then aggregated to the capital-occupation level.

defined as follows:

$$\mathbf{w_{ij}} = \mathrm{TF_{ij}} \cdot \mathrm{IDF_i}\,, \qquad \mathrm{TF_{ij}} = \frac{f_{\mathbf{ij}}}{\sum_i f_{\mathbf{ij}}}\,, \qquad \mathrm{IDF_i} = \log\left(\frac{\mathbf{J}}{\sum_{\mathbf{j}} \mathbb{1}\{\mathbf{i} \in \mathbf{j}\}}\right)\,,$$

where $\mathbf{J}$ is the number of total documents. Therefore, $\mathrm{IDF_{ij}}$ increases when the word appears frequently within the document but decreases when it is common across other documents. This transformation helps match two texts that have meaningful common words. The resulting similarity score ranges from 0 to 1 by construction. A score of 0 indicates that there are no shared words, while a score of 1 demonstrates identical texts.

After constructing similarity scores for each capital goods and task, we aggregate the scores to the capital-occupation level. Since various occupations encompass heterogeneous sets of tasks, they have different scores for a given capital good. We compute the unweighted average of these similarity scores across tasks to obtain scores at the capital-occupation level.

Figure 1 shows the distribution of the similarity scores between capital goods and occupations. The distribution is right-skewed, indicating that most capital-occupation pairs do not have many overlapping words. Nonetheless, some capitals have descriptions closely related to the task descriptions of occupations. For example, in Figure 1, the glass cutter has one of the highest similarity scores with glaziers but a low similarity score with craft artists. Based on these scores, a capital good is considered task-similar to the occupation if the similarity exceeds the 90th percentile; all other capital goods are classified as task-dissimilar.[11]

We follow the imputation procedure of Caunedo et al. (2023) to calculate a quantity index of capital bundles at the occupation and industry level for similar and dissimilar capital. The stock of each category is prorated with an intensity-weighted number of workers in each occupation. Then, an index of capital bundles is measured as a chained index from the base year, 1980, which grows at the weighted average of growth rates across NIPA capital categories using expenditure weights. The user costs of the capital bundles are derived from a series of user costs by capital goods with the zero profit condition in Equation (1). For details on the imputation process, see Appendix B.

Table 1 shows the intensity of capital, defined as the average capital stock per worker among various groups of occupations. Panels A and B sort occupations based on their abstract and routine scores from Autor and Dorn (2013). The intensity of task-similar capital was the highest for non-abstract occupations in the first quintile in both 1980 and 2015, whereas the intensity of task-dissimilar capital was almost flat across the abstract score of occupations in 1980 but became the highest for abstract occupations in the fifth quintile in 2015. In Panel B, routine occupations in the fifth quintile had the fastest growth in task-similar capital between 1980 and 2015. In contrast, their growth of task-dissimilar capital was relatively less pronounced than in non-routine occupations.

In Panel C, occupations are sorted by their labor shares in occupational expenditures on capital and labor in 1980. Low labor shares implied a lower intensity of similar capital but a higher intensity of dissimilar capital in 1980. In 2015, low labor shares were asso-

[11] We try different thresholds and show the estimation results in Appendix D.

Table 1: Capital Intensity across Occupation Groups

	Similar			Dissimilar		
	Low	Middle	High	Low	Middle	High
	Panel A. Across Abstract Score					
1980	17.59	10.01	3.23	27.78	29.25	26.13
2015	36.26	19.56	8.60	73.90	91.54	145.48
	Panel B. Across Routine Score					
1980	2.80	14.53	5.77	8.42	34.20	32.51
2015	4.40	28.18	17.36	39.45	118.25	100.43
	Panel C. Across Labor Share in 1980					
1980	7.94	12.65	6.76	51.85	25.29	8.42
2015	25.74	22.96	9.82	190.40	81.98	30.40
	Panel D. Across Wage in 1980					
1980	9.94	9.09	14.46	24.62	29.15	29.89
2015	17.06	20.24	27.44	52.45	96.68	149.34

Notes. This table presents the capital intensity for task-similar and task-dissimilar capital between occupations segmented into three groups. Capital intensity is defined as the average capital stock per employee, with values expressed in thousands of 2012 dollars. Panel A sorts occupations by their abstract scores, Panel B by routine scores, Panel C by the occupational labor share in 1980, and Panel D by the wage level in 1980. The columns labeled Low and High correspond to occupations in the first and fifth quintiles, respectively, while those labeled Middle encompass occupations within the second to fourth quintiles.

ciated with a higher intensity of similar capital. In relative terms, occupations with high labor shares experienced a more rapid growth of dissimilar capital than similar capital. In contrast, occupations with low labor shares had more balanced growth between similar and dissimilar capital. When occupations are sorted by their wage level in 1980 in Panel D, high-wage occupations in the fifth quintile had the fastest growth of dissimilar capital, whereas the growth of similar capital was comparable across wage groups. As a result, the distribution of dissimilar capital was more dispersed between wage groups in 2015 than in 1980.

In summary, between 1980 and 2015, the intensity of task-similar capital grew faster for non-abstract and routine occupations with low labor shares. In terms of task-dissimilar capital, the intensity increased more rapidly for abstract and non-routine occupations with high wages.

3.3 Measuring Capital-Embodied Innovation

Capital-embodied innovation is measured by matching patents with capital goods. To do so, we calculate the text similarities between patents and capital goods, following a procedure similar to Section 3.2. A patent is assigned to a capital good if the similarity score of its title and abstract to the Wikipedia description of the capital good exceeds the 90th percentile across patent-capital pairs. Some patents may not be relevant to any of the capital goods, and others may be relevant to many. Therefore, we allow patents to be matched to multiple or none of the capital goods based on the similarity score. We allow a single patent to be assigned to, at most, five capital goods. A patent linked to multiple capital goods is weighted by the inverse number of the matched goods. After this procedure, 27% of the patents are matched with at least one capital good.[12] We count the cumulative number of patents associated with each good in the UNSPSC from 1970.

The number of patents across capital goods is then aggregated at the occupation level. Note that each occupation uses two types of capital goods: task-similar and task-dissimilar. Based on the crosswalk between the UNSPSC and NIPA categories from Caunedo et al. (2023), the number of patents at the eight-digit UNSPSC level is averaged within each NIPA capital category for each occupation and capital group. The occupation-level CEI is measured by the average log difference in the number of patents across the NIPA categories, each weighted by the average capital expenditure share between 1980 and 2015, as in Equation (3).[13] By taking averages across capital goods and then capital categories, our CEI measures do not reflect the variety of capital goods within capital categories and the quantity of capital categories within a capital type.

Table 2 presents the summary statistics for the average number of patents at the occupation-industry level.[14] The number of patents has increased over time but at dif-

[12]Table A2 shows the share of patents that matched at least one capital good across patent classes and periods. Example 1 in Appendix A.1 displays a sample matching between patent and capital goods.

[13]In our context, taking the log difference in the number of patents removes time-invariant components of measurement errors associated with the text-matching procedure. For example, if Wikipedia articles about lasers are easier to match than those about computers and the errors are multiplicatively separable and constant over time, log-differencing the number of patents cancels out the errors.

[14]If an occupation does not have any task-similar good, the number of patents is zero for task-similar capital.

Table 2: Summary Statistics of Patents Matched with Capital Goods

	Similar		Dissimilar		
	Mean	SD.	Mean	SD.	N
1970 – 1980	44.4	97.5	45.8	71.2	15,902
1980 – 1990	93.5	185.6	96.0	133.3	15,902
1990 – 2000	137.4	271.1	171.5	225.3	15,902
2000 – 2015	428.5	840.5	558.1	667.6	15,902

Notes. This table displays the summary statistics of patents matched with task-dissimilar and task-similar capital aggregated at the occupation-industry level. We take the average number of patents, weighted by the share of capital expenditure in each period.

Table 3: CEI Measure across Occupation Groups

Similar			Dissimilar		
Low	Middle	High	Low	Middle	High
Panel A. Across Abstract Score					
2.40	1.70	1.77	2.90	3.58	3.78
Panel B. Across Routine Score					
0.81	2.34	1.52	3.59	3.49	3.30
Panel C. Across Labor Share in 1980					
2.66	1.80	1.04	3.11	3.47	3.95
Panel D. Across Wage in 1980					
2.17	1.84	1.57	3.05	3.45	4.01

Notes. This table presents the employment-weighted averages of CEI across three bins of occupations. Panel A categorizes occupations based on their average wages in 1980, Panel B uses the abstract score, and Panel C employs the routine score. The CEI is defined in Equation (3). The columns labeled Low and High represent the occupations in the first and fifth quintiles, respectively, while those labeled Middle cover occupations within the second to fourth quintiles.

ferent rates across occupations. Initially, the number of patents was comparable between similar and dissimilar capital. However, task-dissimilar capital experienced faster growth in patents than task-similar capital, suggesting that more patents are made on capital goods used as task-dissimilar capital.

Table 3 displays CEI for various occupational groups. Panel A sorts occupations by their abstract task scores, showing that CEI-s is the highest in the first quintile, while CEI-d is the highest in the fifth quintile. These patterns suggest that abstract occupations are disproportionately more affected by innovation in dissimilar capital and less by innovation in similar capital. When occupations are sorted by routine task scores in Panel B, non-routine occupations in the first quintile have the highest CEI-d, but the dispersion is not as strong as when occupations are sorted by abstract task scores. Furthermore, CEI-s is not monotone in routine task scores.

In Panels C and D, occupations are grouped by their labor share and wage levels in 1980, respectively. CEI-s is higher for occupations with low labor shares, whereas CEI-d is higher for occupations with high labor shares. Finally, low-wage occupations have relatively higher CEI-s, while high-wage occupations have relatively higher CEI-d. To summarize, CEI-d was biased towards abstract and non-routine occupations with high labor share and wages. CEI-s is higher for non-abstract and low-wage occupations with low labor share. These findings align with the heterogeneous growth of capital intensity in Section 3.2. Our results confirm that occupations are heterogeneously exposed to innovations in magnitude and composition, which is consistent with the hypothesis of Autor et al. (2003), who attribute heterogeneous changes in occupational labor demand to computer technology becoming popular over time.

Since some tools are not listed on the Wikipedia page and are therefore excluded from our calculation, the CEI for certain capital goods with higher missing rates may be underestimated. To address this issue, we reweight the tools by the inverse of their Wikipedia matching rate and report that the results remain quantitatively similar in Appendix Table A2.

4 Estimation

4.1 Strategy

Estimation of γ_p, the coefficient of CEI on capital user costs, is straightforward. We apply Equation (2) to estimate γ_p.

$$\Delta \log r_{jio} = \gamma_p \Delta \log P_{jio} + \omega_{jio}. \tag{15}$$

Unless otherwise noted, Δ denotes an operator that takes the difference between 1980 and 2015. r_{jio} represents the user cost of capital for capital type j in industry i and occupation o, as defined in Equation (1). $\Delta \log P_{jio}$ is defined in Equation (3) as the average growth rate of patents matched to each capital category $n \in \mathbb{N}_{jo}$, weighted by the average capital expenditure share in 1980 and 2015.[15]

Next, we use the first-order conditions of the cost minimization in Section 2.3 for estimation. Specifically, Equation (10) is used to estimate the elasticities of substitution for the inner CES composite, for the outer CES composite, and across occupational services. Combining Equations (7) and (8) with (10) gives

$$\begin{aligned} \Delta \log l_{io} = {} & \gamma_s \Delta \log P_{sio} + \gamma_d \Delta \log P_{dio} + \kappa_a \Delta \log \left(w_o\right) \\ & + \kappa_s \Delta \log \left(1 + \frac{r_{sio} k_{sio}}{w_o l_{io}}\right) \\ & + \kappa_d \Delta \log \left(1 + \left(1 + \frac{r_{sio} k_{sio}}{w_o l_{io}}\right)^{-1} \times \frac{r_{dio} k_{dio}}{w_o l_{io}}\right) + X_{io}' \beta + \nu_{io}. \end{aligned} \tag{16}$$

In this equation, $\kappa_s = \frac{\sigma - \rho_s}{\rho_s - 1}$, $\kappa_d \equiv \frac{\sigma - \rho_d}{\rho_d - 1}$, and $\kappa_a \equiv -\sigma$. ν_{io} is the residual demand components, and X_{io} is a set of control variables, which are described below.

Several endogeneity concerns arise for the estimation of this equation. First, CEI measures $\Delta \log P_{sio}$ and $\Delta \log P_{dio}$ may be correlated with ν_{io} if innovation activities respond

[15] The regression is conducted at the occupation-industry level, where patents and capital goods are aggregated. We adopt this specification because the occupation-industry level is the primary unit of analysis in our study. Alternatively, we run the regression at the capital good level and confirm that the regression coefficients are quantitatively similar in Section 4.2.

to occupational demand shocks. For example, an increase in supply for certain occupational services could incentivize firms to make more innovations in related capital goods. Second, since w_o is jointly determined with l_{io}, we need an exogenous shifter for labor supply. Lastly, the capital-to-labor income ratios for task-similar and task-dissimilar capital may also be correlated with ν_{io} if investment decisions are affected by the residual demand shocks.

To address these concerns, we introduce five instrumental variables: academic publications related to task-similar and task-dissimilar capital, immigration shocks from Latin American countries, and changes in import value for each type of capital. Our identification assumption is that these instrumental variables are correlated with the main independent variables but not with the residual demand shocks, conditional on controls. Appendix C details the variations and relevance of the instruments.

We construct shift-share instruments for CEI that capture heterogeneous knowledge spillovers from academic publications to patents. For example, innovation in the computer sector is based on knowledge produced in the field of electronic engineering. An increase in the number of papers in electronic engineering is positively correlated with innovation in the computer sector but unlikely to be correlated with demand shocks for IT workers. This IV approach is similar to that of Aghion et al. (2019) and Berkes et al. (2022), who use patent citations across regions and countries, respectively, to construct the share and use patent growth as a shifter. In contrast, we exploit citations from patents to academic publications to construct the share and use the growth rate of publications in European countries as the shifter to obtain more exogenous variation.

To measure the diffusion of knowledge from academic publications to patents, we use citation data from patents to academic publications following the literature (Marx and Fuegi, 2020; Arora et al., 2021).[16] We first construct the upstreamness of an academic field f to the patent class p by using citations made from 1970 to 1980.[17] The upstreamness

[16] A large number of citations from patents within a technology class to papers in a specific academic field suggests that the academic field serves as an upstream source of knowledge for that technology class. Marx and Fuegi (2020) show that 17.6% of USPTO patents cite at least one academic paper, with an average of two academic citations per patent.

[17] For the academic field, the Web of Science Field is used, encompassing 251 distinct fields. For patents,

υ_{pf} is calculated as below:

$$\upsilon_{pf} = \frac{\mathcal{C}_{pf}}{\sum_f \mathcal{C}_{pf}} ,$$

where $\mathcal{C}_{pf}$ is the number of citations from patent class p to academic field f. The left panel of Appendix Figure A1 plots the variation of citation share over patent classes.

Next, we use this share to construct the exposure measure of the capital bundle k_{jio} to the academic field f. The upstream measure υ_{pf} is multiplied by $s^{Pat.}_{jiop}$, the average share of patent class p in capital bundle k_{jio}, weighted by the expenditure-adjusted number of patents from 1970 to 1980.[18]

$$z^{Pub.}_{jio} = \sum_p s^{Pat.}_{jiop} \sum_f \upsilon_{pf} \Delta \log(\mathcal{P}_f) , \tag{17}$$

where $\mathcal{P}_f$ is the number of publications in field f. Papers affiliated only with European institutions are counted to avoid the potential bias from US patenting firms that also engage in academic projects, which could be correlated with residual demand shocks. The instrument increases if the academic fields relevant to the capital bundle experience faster growth in publication.

The immigration instrument uses heterogeneous shares of immigrants from Latin American countries at the occupation and industry level interacted with the growth of Latin American immigration. Immigrants from Latin American countries likely possess comparative advantages that differ from those of US-born workers, influencing occupational labor supply in distinct ways. For each occupation, the heterogeneous exposure to immigration shocks is computed based on the share of Latin American workers in 1980. The Bartik immigration shock for occupation o is defined by the following equation.

$$z^{Latin}_{io} = \sum_c s^{Latin}_{co} \Delta \log \left(\mathbf{L}_c - \mathbf{l}_{co} \right) . \tag{18}$$

In the equation, s^{Latin}_{co} is the share of workers from Latin American country c in occupation

three-digit IPC patent classes, comprising 387 classes, are used.

[18] Appendix C.1 provides the formal definition of $s^{Pat.}_{jiop}$.

o in 1980. $\mathbf{l}_{co}$ is the number of workers in occupation o from country c, and $\mathbf{L}_c$ is the total number of immigrants. Workers in occupation o are subtracted from calculating the supply shock to rule out the effect of occupation-level shocks that led to more immigrants from Latin America.

Lastly, we use capital import data to construct a shifter for capital expenditure. Specifically, we calculate the log change in import value from 1980 to 2015 for each capital category n. Then, the import shock is defined as below:

$$z_{jio}^{Import} = \sum_{n \in \mathbb{N}_{jio}} s_{jion}^{Import} \Delta \log m_n \,, \tag{19}$$

where s_{jion}^{Import} represents the capital expenditure share of capital good n for capital bundle k_{jio} in 1980. m_n is the value of imports. We expect that an increase in the value of imports would be positively associated with an increase in capital stock from 1980 to 2015. For example, if a capital bundle relied more on sewing machines and its import increased significantly due to cheaper supplies from other countries, the capital expenditure $r_{jio}k_{jio}$ would also increase. This global increase in commodity-level trades is plausibly exogenous to the demand shock at the occupation level between 1980 and 2015.

Controls include the task-offshorability index at the occupation level from Autor and Dorn (2013) and the initial levels of log wage bill. The offshorability index controls demand changes for occupational services related to international trade. We also formulate a Bartik-style control for occupational services with industry-level employment growth interacting with industry composition at the occupation level. Lastly, one-digit occupation fixed effects are added in the regression, which implies that the regression compares occupations in the same categories, such as managerial occupations or professional specialty occupations. 123 occupations that do not have any task-similar capital are excluded from the estimation.

We exploit the variations at the industry-by-occupation level for two reasons. First, industries have different capital composition. For example, the manufacturing industry works more with machinery than the service industry. Thus, our CEI measure of opera-

Table 4: First-Stage Results

	Wage	Patent-s	Patent-d	Capital-s	Capital-d
Immigration	-1.42	0.00	0.12	0.02	0.00
	(0.02)	(0.02)	(0.03)	(0.00)	(0.00)
Publication-s	-0.21	1.80	0.08	0.02	0.02
	(0.07)	(0.08)	(0.11)	(0.01)	(0.01)
Publication-d	-0.48	0.49	1.67	-0.07	0.05
	(0.07)	(0.08)	(0.11)	(0.01)	(0.01)
Import-s	0.42	0.54	-0.63	0.11	0.03
	(0.04)	(0.05)	(0.07)	(0.00)	(0.00)
Import-d	0.28	0.04	0.32	-0.05	0.03
	(0.05)	(0.05)	(0.08)	(0.01)	(0.01)
N	7,645	7,645	7,645	7,645	7,645
F-statistics	900.15	122.70	131.75	139.06	74.67

Notes. This table presents the results of the first-stage regression in Equation (16). Standard errors are in parentheses. Immigration is the instrumental variable defined in Equation (18). Import-d and Import-s represent import shocks for task-similar and task-dissimilar capital, respectively, as characterized in Equation (19). Publication-d and Publication-s represent publication shocks for task-similar and task-dissimilar capital, respectively, as defined in Equation (17). Wage is $\Delta \log w_o$. Patent-s and Patent-d are $\Delta \log P_{sio}$, $\Delta \log P_{dio}$, respectively. Capital-s is $\left(1 + \frac{r_{sio}k_{sio}}{w_o l_{io}}\right)$, and capital-d is $\Delta \log \left(1 + \left(1 + \frac{r_{sio}k_{sio}}{w_o l_{io}}\right)^{-1} \times \frac{r_{dio}k_{dio}}{w_o l_{io}}\right)$.

tions researchers gives more weight to machine-related patents for the manufacturing industry than for the service industry. In Appendix E, we decompose the variations of CEI measures and their instruments between occupations and industries. The results suggest that more than 20% of the variations in CEI-d are the result of variations within occupations. Second, occupational composition varies significantly between industries. Thus, it is important to control heterogeneous trends between industries when quantifying the effect of CEI on labor demand.

4.2 Results

Table 4 presents the results of the first stage. All instrumental variables display signs with the corresponding variables consistent with the prior. Specifically, immigration shocks are negatively correlated with wage changes, and publication shocks for each type of capital are positively correlated with the patent measure. In addition, import shocks

Table 5: Parameter Estimates

	ρ_s	ρ_d	σ	γ_p	γ_s	γ_d
Estimate	3.02	1.55	2.27	-0.41	-0.59	0.86
SE	(0.97)	(0.24)	(0.24)	(0.01)	(0.18)	(0.31)

Notes. This table shows the estimates and standard errors of the regression equations (15) and (16). ρ_s (ρ_d) is the elasticity of substitution between task-similar (task-dissimilar) capital and labor. σ is the elasticity of substitution between different occupational services. γ_p is the coefficient of CEI on user costs of capital. γ_s (γ_d) is the coefficient of CEI-s (-d) on occupational service demand shifter.

for each type of capital are positively associated with the corresponding capital income ratios. Each regression exhibits high F-statistics, and the Cragg-Donald Wald F-statistic value is 14.06, indicating that the instrumental variables are strong in the first stage.

Table 5 shows the estimation results.[19] The elasticity of substitution between task-similar capital and labor is estimated at 3.0, whereas the elasticity between labor and task-dissimilar capital is 1.6. Our results are modestly higher than the estimates in Caunedo et al. (2023), who assume a single elasticity of substitution between capital and labor for each occupation. Using time-series variations in birth rates and the supply of educated workers to construct the occupation-level supply shifter, they find that the elasticity ranges from 0.7 to 2.2. The estimation in this paper deals with long-term adjustments in the labor market over three decades and uses cross-sectional variations. We also construct a labor supply shifter with cross-sectional exposure to immigration from Latin America. Our marginally higher estimates are likely to result from dealing with a longer time horizon and cross-sectional variations for estimation.

The substitution elasticity across occupational labor inputs, σ, is also estimated at a value higher than that of the literature. In Caunedo et al. (2023), the value is calibrated at 1.3. This paper estimates the value of 2.3. This value is larger than the estimates by Lee and Shin (2017), 0.7, and Burstein et al. (2019), 2. The higher estimate of the elasticity *between* occupational services in this paper is made with the longer time horizon for

[19]Table A5 in Appendix D.2 reports estimation results when we impose an alternative nesting in the occupational production function. Table A6 in Appendix D.3 presents the results where tools are weighted by the inverse of their finding rates in Wikipedia for CEI estimation.

adjustments. In addition, these papers use more aggregated levels of occupation codes. Caunedo et al. (2023) and Lee and Shin (2017) report results with 11 occupations. Burstein et al. (2019) have variations of 30 occupations. We have 291 occupations distinguished by three-digit occupation codes from the 1990 Census.

The estimate for σ is smaller than ρ_s but larger than ρ_d. Due to the high standard error of ρ_s, it is hard to tell whether ρ_s is significantly greater than ρ_d or σ. However, ρ_d is less than σ at the significance level of 1%. As discussed in Section 2.3, these values imply that the scale effect of CEI increasing the demand for occupational services is smaller than the substitution effect between labor and capital for task-similar capital, but the reverse is true for task-dissimilar capital. As a result, a lower user cost of task-similar capital reduces the relative labor demand. Conversely, a lower user cost of task-dissimilar capital raises the relative labor demand.

Table 5 also presents the estimates for the coefficient of CEI measures on the user costs of capital and demand shifters for occupational services. The estimate for γ_p is negative and large in magnitude.[20] An increase in CEI by 1% reduces the user cost of capital by 0.41%. Along with the estimates of elasticities, this negative coefficient estimate implies that CEI-s leads to stronger substitution with capital, decreasing occupational labor demand. In contrast, CEI-d stimulates the demand for occupational services more, which dominates the substitution effect and raises labor demand. The estimate for γ_s is negative, and the estimate for γ_d is positive. These results indicate that CEI-s decreases the demand for occupational services, whereas CEI-d raises it, even after considering their effects on user costs. These effects are also quantitatively important. An increase in CEI-s by 1% reduces occupational service demand by 0.59%, and an increase in CEI-d by 1% increases occupational service demand by 0.86%.

In counterfactual exercises, we consider two values for the elasticity of occupational labor supply: $\eta = 0.3$ or $\eta = 1$. Caunedo et al. (2023) calibrate $\eta = 0.3$ at the yearly frequency and with coarser occupational codes. Since we consider labor supply adjust-

[20] Alternatively, we run the regression at the NIPA capital good and industry level. Then, the coefficient of CEI on the user costs of a capital good is estimated to be -0.47 with a standard deviation of 0.06.

ments over longer than three decades with more detailed occupational codes, the supply elasticity calibrated at 0.3 is likely to be a lower bound.[21] To capture the possibility that labor supply is more elastic to wage changes, counterfactual equilibrium with $\eta = 1$ is also derived in Section 5.

5 Counterfactuals

The counterfactual exercise aims to address the following question: What happens to the labor market without CEI that is heterogeneous between occupations and industries? To address this question, we calculate counterfactual equilibria where CEI measures are back to their levels in 1980, with other demand and supply shocks unchanged.

5.1 CEI and Task-Biased Labor Market Changes

We test what task-biased changes in the labor market would look like without CEI between 1980 and 2015 and thereby see if CEI constitutes task-biased technical changes in Autor et al. (2003). The task bias of labor market changes is measured in the following auxiliary regression that relates the log differences in wage and employment between 1980 and 2015 to abstract and routine task scores at the occupation level from Autor and Dorn (2013).

$$\text{Change}_o = \theta_0 + \theta_1 \text{Task Score}_o + \varepsilon_o. \tag{20}$$

Each occupation is weighted by its employment in 1980. The OLS estimates for θ_1 summarize the correlation between abstract and routine task scores with cross-sectional changes in wage and employment at the occupation level and thus are used as a measure of task bias of labor market changes in 1980-2015.

Table 6 shows the estimates for the regression coefficients. During this period, if an

[21] We regress the following supply equation: $\Delta \log L_o = \psi + \eta \Delta \log w_o + \epsilon_o$ with CEI and import shocks as demand instruments at the occupation level. Then η is estimated at 2.1. Thus, we take a value between the two numbers, 1, as our benchmark.

occupation has a one-standard-deviation higher score on abstract tasks, it has 0.17 and 0.44 standard deviations higher employment and wage growth rates, respectively.

In Panel A of Table 6, two counterfactual equilibria are computed for $\eta = 0.3$ and $\eta = 1$ when the patent measures return to their levels in 1980. The biases in employment and wage are smaller for both abstract and routine task scores in the absence of CEI. Our results suggest that CEI explains 18–59% of abstract-biased employment changes and almost all wage growth. Likewise, CEI could account for 8–27% and 70–79% of the bias against routine occupations in wage and employment growth, respectively.

The effect on the wage is larger than the effect on employment, particularly when $\eta = 0.3$. This is due to the calibrated value of the supply elasticity η being too low to generate large responses in employment. Between 1980 and 2015, employment changes were more dispersed, showing a standard deviation of 0.75, compared to wage changes, which had a standard deviation of 0.22. Thus, with the values of supply elasticity considered, most employment changes result from non-wage supply shifters, ϵ_o. With $\eta = 1$, the employment responses become larger, while the wage responses diminish. Since $\eta = 0.3$ likely represents the lower bound, $\eta = 1$ is assumed for the counterfactual exercises below.

Panel B shows the counterfactual results, where only one type of CEI is fixed to the level in 1980. CEI-s and CEI-d have qualitatively similar effects on the task bias of employment and wage changes. CEI-s is higher for abstract and non-routine occupations, whereas CEI-d is higher for abstract ones, as shown in Table 3. Consequently, the negative effect of CEI-s yields results analogous to the positive effect of CEI-d on occupational service demand. Given the comparable magnitudes of the coefficient estimates in Table 5, CEI-d and CEI-s have similar quantitative effects.

Panel C decomposes the two channels in which CEI affects occupational labor demand. Almost all employment responses operate through demand shifters. Lower user costs of capital from CEI have limited effects on employment. This is because the estimated gaps between ρ_s, ρ_d, and σ are small. In Equation (10), the exponents of $\tilde{\Theta}_{io}$ and $\tilde{y}_{io}$ are $(\rho_d - \rho_s)/(\rho_s \times \rho_d) = -0.31$ and $-1/\sigma + 1/\rho_d = 0.20$, respectively. Together, a 1% in-

Table 6: Counterfactual: Task-Biased Labor Market Changes

	Abstract		Routine	
	Employment	Wage	Employment	Wage
Actual Change	0.17	0.44	-0.26	-0.33
	Panel A. Varying Supply Elasticity			
Without CEI (η=0.3)	0.14	0.02	-0.24	-0.07
Without CEI (η=1)	0.07	0.05	-0.19	-0.10
	Panel B. Similar vs. Dissimilar CEI			
Without CEI-s	0.13	0.19	-0.22	-0.18
Without CEI-d	0.14	0.20	-0.23	-0.17
	Panel C. Different Channels			
Through User Costs	0.19	0.44	-0.26	-0.33
Through Task Demand	0.08	0.06	-0.19	-0.10

Notes. This table shows the actual and counterfactual regression coefficients of the wage and employment growth rates at the occupation level in the occupational task scores of Autor and Dorn (2013) as in Equation (20). The counterfactual equilibrium fixes the CEI measures at their levels in 1980. The rows $\eta = 0.3$ and $\eta = 1$ in Panel A set the elasticity of occupational labor supply at 0.3 and 1, respectively. Panel B fixes patent measures of either similar or dissimilar capital at the 1980 level separately. Panel C fixes patent measures to the 1980 level only when calculating changes in user costs and occupational demand, respectively. Panels B and C assume $\eta = 1$.

crease in $\tilde{\Theta}_{io}$ reduces the occupational labor demand in Equation (10) by less than 0.11%, and a 1% increase in $\tilde{y}_{io}$ increases it only by 0.20%. Moreover, ρ_d and ρ_s are close to one, weakening the effect of CEI-s and CEI-d on $\tilde{\Theta}_{io}$ and $\tilde{y}_{io}$ in Equations (11) and (12). Thus, the user cost channel has a minor role in generating significant employment responses. In contrast, the estimated coefficients of CEI on demand shifter are large.

5.2 CEI and Declines in Labor Share

We use our framework to examine how CEI has contributed to the decline in labor share in the United States since the late 1970s (Elsby et al., 2013; Karabarbounis and Neiman, 2014). We first express the aggregate labor share as the expenditure-weighted average of the labor share at the occupation-industry level.

$$LS = \sum s_{io}^{Input} \frac{w_o l_{io}}{w_o l_{io} + r_{sio} k_{sio} + r_{dio} l_{dio}}. \tag{21}$$

In this equation, s_{io}^{Input} represents the share of occupation o and industry i in the total expenditure on production inputs. Using this equation, we decompose the change in the aggregate labor share into two components: the within-occupation margin, which captures the change in labor share with fixed income shares, and the between-occupation margin, which reflects the change in labor share with fixed within-occupation labor shares.

In our framework, lower user costs of capital associated with CEI change aggregate labor share within and between occupations. First, because the elasticities of substitution of labor with task-similar and task-dissimilar capital are greater than one, lower user costs decrease labor shares within occupations and industries. Second, as the elasticity of substitution between occupational services is also greater than one, a lower user cost increases the input cost share of an occupation. Since occupations have different labor shares, changes in input cost shares affect the aggregate labor share.

Changes in occupational demand shifters also alter the aggregate labor share within and between occupations. Because the supply elasticity is greater for capital than labor, an increase in occupational service demand increases wage relative to user costs and reduces within-occupation labor shares. At the same time, since CEI varies at the occupation level, it directly shifts input cost shares between occupations.

Table 7 summarizes the actual and counterfactual changes in labor share. The data show that the aggregate labor share dropped from 81% to 71% between 1980 and 2015. These numbers are higher than the labor share in national accounts because they do not include profits and capital expenditures on structures. Both within-occupation and across-occupation margins contribute to lower labor share, but within-occupation margin is quantitatively more important in driving the changes.

Panel A indicates that, without CEI, the aggregate labor share falls by 1.1–1.4%, 86–89% smaller than the actual change. CEI especially contributes to larger declines in

Table 7: Counterfactual: Declines in Labor Share

	Within	Between	All
Actual Change	-8.64	-1.24	-9.87
	Panel A. Varying Supply Elasticity		
Without CEI (η=0.3)	-0.06	-1.04	-1.09
Without CEI (η=1)	-0.18	-1.24	-1.42
	Panel B. Similar vs. Dissimilar CEI		
Without CEI-s	-6.82	-0.80	-7.62
Without CEI-d	-3.10	-0.96	-4.06
	Panel C. Different Channels		
Through User Cost	-0.42	0.80	0.38
Through Task Demand	-0.46	-1.21	-1.67

Notes. This table shows changes in labor share between 1980 and 2015 in actual data and counterfactual equilibria. The counterfactual equilibrium fixes the CEI measures at their levels in 1980. The column Within indicates changes in labor share from changes with fixed input income share between occupations (within-occupation), and the column Between shows changes in labor share with fixed within-occupation labor shares (across-occupation). The rows $\eta = 0.3$ and $\eta = 1$ in Panel A set the elasticity of occupational labor supply at 0.3 and 1, respectively. Panel B fixes patent measures of either similar or dissimilar capital to the 1980 level separately. Panel C fixes patent measures to the 1980 level only when calculating changes in user costs and occupational demand, respectively. Panels B and C assume $\eta = 1$.

within-occupation labor shares but barely changes between-occupation labor shares. In Panel B, CEI-d has a larger effect on within-occupation labor share than CEI-s because CEI-d is higher than CEI-s in Table 3. Moreover, some occupations do not have task-similar capital. Consequently, although the elasticity of substitution of task-dissimilar capital with labor is closer to one than that of task-similar capital, CEI-d lowers within-occupation labor shares more substantially than CEI-d. Due to the non-linearity of the model, the sum of their separate effects in Panel B is different from the combined effects of CEI-s and CEI-d in Panel A, particularly in the between-occupation margin.

In Panel C, both user cost and task demand channels significantly reduce within-occupation labor shares because of elastic substitution toward capital and inelastic labor supply. Meanwhile, changes in user costs also contribute to decreases in labor share between occupations by changing the share of input costs between occupation and industry. Occupations with high labor shares exhibit higher CEI-d but lower CEI-s, resulting in opposing effects from changes in user costs. In aggregate, the effect of CEI-s dominates, thereby reducing the share of high-labor-share occupations. Higher CEI-d and lower CEI-s of these occupations shift up the task demand and income share of these occupations. However, the inelastic labor supply increases their wages, lowers their input cost shares, and dampens the demand shifter effect. Consequently, the task demand channel also has a significant effect on labor share within these occupations.

5.3 CEI and Labor Market Polarization

Table 8: Counterfactual Polarization

	Employment			Wage		
	Low	Middle	High	Low	Middle	High
Actual Change	0.17	-0.58	0.41	-0.03	-0.54	0.57
	Panel A. Varying Supply Elasticity					
Without CEI (η=0.3)	0.21	-0.59	0.38	0.21	-0.33	0.12
Without CEI (η=1)	0.26	-0.56	0.30	0.20	-0.36	0.16
	Panel B. Similar vs. Dissimilar CEI					
Without CEI-s	0.28	-0.63	0.35	0.19	-0.51	0.32
Without CEI-d	0.24	-0.58	0.34	0.07	-0.33	0.26
	Panel C. Different Channels					
Without Δ User Costs	0.16	-0.57	0.41	-0.04	-0.53	0.57
Without Δ Occ. Demand	0.24	-0.53	0.29	0.19	-0.36	0.17

Notes. This table shows the actual and counterfactual growth rates of wage and employment when occupations are grouped by their wages in 1980. The counterfactual equilibrium fixes the CEI measures at their levels in 1980. The wage and employment changes are standardized to have a zero mean and a standard deviation of one in each case. Columns labeled Low and High denote occupations with 1980 wage levels in the first and fourth quintiles, respectively. Columns labeled Middle denote occupations between the two quintiles. The rows $\eta = 0.3$ and $\eta = 1$ in Panel A set the elasticity of occupational labor supply at 0.3 and 1, respectively. Panel B fixes patent measures of either similar or dissimilar capital to the 1980 level separately. Panel C fixes patent measures to the 1980 level only when calculating changes in user costs and occupational demand, respectively. Panels B and C assume $\eta = 1$.

Lastly, we study how CEI is related to polarization in the labor market after the 1980s. The first row of Table 8 summarizes wage and employment changes between 1980 and 2015 for three occupation groups by their residual wages in 1980. As in Autor and Dorn (2013), employment and wage changes at the occupation level take a U-shaped form over the wage level in 1980. High-wage occupations in the fifth quintile have 0.41 and 0.57 standard deviations higher employment and wage growth rates than the average occupations, respectively, and low-wage occupations also exhibit higher cross-sectional wage and employment growth rates than the average. The middle-wage ones, on the contrary, experience lower growth in terms of both employment and wage; the employment growth rate is lower by 0.58 standard deviations than the average, and the wage growth rate is lower by 0.54 standard deviations than the average.

Panel A of Table 8 shows the counterfactual employment and wage growth when the patent measure, Patent_{jon}, is fixed at the level in 1980. Under the counterfactual equilibrium without CEI, wage and employment increase less for high-wage occupations. This result comes from high-wage occupations with higher CEI-d and lower CEI-s than others. The reduction in high-wage employment is driven more by increases in low-wage occupations than by middle-wage ones. CEI has a smaller impact on demand for middle-wage occupations because low-wage occupations have a higher level of CEI-s and a lower level of CEI-d than middle-wage occupations, as shown in Table 3.

Panel B of Table 8 shows the polarization measures from counterfactual equilibria when only one type of capital has the patent measure fixed to 1980. Both CEI-s and CEI-d contribute to the growth of high-wage occupations, and the effects are quantitatively similar to each other. The last panel summarizes the counterfactual results when the impact of CEI is active for only the user costs of capital or the demand shifter in the production function at each time. Similarly to the task bias exercise, the vast majority of wage and employment changes arise from the demand-shifter channel. This is also because the estimated gaps between elasticities of substitutions are estimated to be too small to have substantial employment effects.[22]

[22] Appendix G displays the effect of CEI on user costs of capital and how user costs alone can generate labor market changes. Appendix H presents the results when only computers are included in the CEI measure. In summary, the results are qualitatively consistent with the baseline but with a much smaller magnitude.

6 Conclusion

This paper develops a measure of capital-embodied innovation (CEI) by matching patents with descriptions of capital goods from Wikipedia using text-based analysis. We then use this measure to study the impact of technological factors on labor market trends. Differences in the use of capital goods at the occupation level provide effective cross-sectional variations to identify the impact of CEI.

This paper also proposes a crucial factor that determines the relationship between technological changes in capital and labor demand: the similarity between occupational tasks and the functions of capital. If capital functions similarly to occupational tasks, technological changes that reduce the user costs of such capital promote substitution towards capital and decrease labor demand. Conversely, if capital performs functions different from but essential to occupational tasks, technological changes in this type of capital increase the relative labor demand for occupational labor. This distinction implies that the effect of CEI is strongly dependent on the relationship between capital functions and occupational tasks.

CEI explains historical labor market trends, such as task-biased changes, declines in labor share, and job polarization. Counterfactual analysis shows that CEI accounts for a significant portion of employment growth, particularly in abstract and high-wage occupations, due to the higher CEI for their task-dissimilar capital and lower CEI for their task-similar capital. CEI also contributed to historical declines in labor share by reducing the user costs of capital and within-occupation labor shares. In addition, CEI reallocates expenditures to occupations with low labor shares, thus decreasing the aggregate labor share.

Using CEI measures from patents, technological factors can be distinguished from others, such as trade and outsourcing. Innovations have shaped biased trends in labor market demand, indicating that innovation policies can create biased labor market changes. Since these policies affect innovations in various types of capital to different extents—and because occupations are exposed to capital differently—innovation policies have heterogeneous consequences between occupations. Therefore, supplementary policies need to target more exposed occupations to reduce structural unemployment and lower inequality in the labor market. The results in this paper highlight the need for ongoing research on the long-term responses of the labor market to innovation policies through CEI.

References

Acemoglu, Daron and Pascual Restrepo, "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 2018, *108* (6), 1488–1542.

_ **and** _ , "Tasks, Automation, and the Rise in US Wage Inequality," *Econometrica*, 2022, *90* (5), 1973–2016.

Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hémous, "Innovation and Top Income Inequality," *The Review of Economic Studies*, 2019, *86* (1), 1–45.

Argente, David, Salomé Baslandze, Douglas Hanley, and Sara Moreira, "Patents to Products: Product Innovation and Firm Dynamics," *Working Paper*, 2023.

Arora, Ashish, Sharon Belenzon, and Lia Sheer, "Knowledge Spillovers and Corporate Investment in Scientific Research," *American Economic Review*, 2021, *111* (3), 871–98.

Autor, David, Caroline Chin, Anna Salomons, and Bryan Seegmiller, "New Frontiers: The Origins and Content of New Work, 1940–2018," *The Quarterly Journal of Economics*, 2024, p. qjae008.

Autor, David H and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 2013, *103* (5), 1553–1597.

_ **, Frank Levy, and Richard J Murnane**, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333.

_ **, Lawrence F Katz, and Melissa S Kearney**, "The Polarization of the US Labor Market," *American Economic Review*, 2006, *96* (2), 189–194.

Berkes, Enrico, Kristina Manysheva, and Martí Mestieri, "Global Innovation Spillovers and Productivity: Evidence from 100 Years of World Patent Data," 2022.

Berlingieri, G, D Lashkari, F Boeri, and J Vogel, "Capital-Skill Complementarity in Firms and in the Aggregate Economy," *Working Paper*, 2022.

Bloom, Nicholas, John Van Reenen, and Heidi Williams, "A Toolkit of Policies to Promote Innovation," *Journal of Economic Perspectives*, 2019, *33* (3), 163–184.

_, Tarek Alexander Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun, "The Diffusion of Disruptive Technologies," *NBER Working Paper*, 2021, (28999).

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel, "Changes in Between-Group Inequality: Computers, Occupations, and International Trade," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 348–400.

Bárany, Zsófia L and Christian Siegel, "Job Polarization and Structural Change," *American Economic Journal: Macroeconomics*, 2018, *10* (1), 57–89.

Caunedo, Julieta, David Jaume, and Elisa Keller, "Occupational Exposure to Capital-Embodied Technical Change," *American Economic Review*, 2023, *113* (6), 1642–1685.

Dierdorff, EC, DW Drewes, and JJ Norton, "O*NET Tools and Technology: A Synopsis of Data Development Procedures," 2006.

Elsby, Michael WL, Bart Hobijn, and Ayşegül Şahin, "The Decline of the US Labor Share," *Brookings papers on economic activity*, 2013, *2013* (2), 1–63.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Megan Schouweiler, and Michael Westberry, "IPUMS CPS: Version 11.0 [dataset]," 2023.

Goldin, Claudia and Lawrence F Katz, *The Race Between Education and Technology*, Harvard University Press, 2008.

Goos, Maarten, Alan Manning, and Anna Salomons, "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, 2014, *104* (8), 2509–26.

— **and** —, "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *The Review of Economics and Statistics*, 2007, *89* (1), 118–133.

Griliches, Zvi, "Capital-Skill Complementarity," *The Review of Economics and Statistics*, 1969, pp. 465–468.

Hémous, David, Morten Olsen, Carlo Zanella, and Antoine Dechezleprêtre, "Induced Automation Innovation: Evidence from Firm-level Patent Data," *Journal of Political Economy*, 2025, *133* (6).

Hornstein, Andreas, Per Krusell, and Giovanni L Violante, "The Effects of Technical Change on Labor Market Inequalities," *Handbook of Economic Growth*, 2005, *1*, 1275–1370.

Jorgenson, Dale W, "Capital Theory and Investment Behavior," *The American Economic Review*, 1963, *53* (2), 247–259.

Karabarbounis, Loukas and Brent Neiman, "The Global Decline of the Labor Share," *The Quarterly journal of economics*, 2014, *129* (1), 61–103.

Keller, Wolfgang and Hâle Utar, "International Trade and Job Polarization: Evidence at the Worker Level," *Journal of International Economics*, 2023, *145*, 103810.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation Over the Long Run," *American Economic Review: Insights*, 2021, *3* (3), 303–20.

Kogan, Leonid, Dimitris Papanikolaou, Lawrence Schmidt, and Bryan Seegmiller, "Technology and Labor Displacement: Evidence from Linking Patents with Worker-Level Data," *NBER Working Paper*, 2023, (w31846).

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante, "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

Lee, Sang Yoon Tim and Yongseok Shin, "Horizontal and Vertical Polarization: Task-Specific Technological Change in a Multi-Sector Economy," *NBER Working Paper*, 2017, (w23283).

Mann, Katja and Lukas Püttmann, "Benign Effects of Automation: New Evidence from Patent Texts," *Review of Economics and Statistics*, 2023, *105* (3), 562–579.

Marx, Matt and Aaron Fuegi, "Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles," *Strategic Management Journal*, 2020, *41* (9), 1572–1594.

Michaels, Guy, Ashwini Natraj, and John Van Reenen, "Has ICT Polarized Skill Demand? Evidence from Eleven Countries over Twenty-Five Years," *Review of Economics and Statistics*, 2014, *96* (1), 60–77.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rogers, and Megan Schouweiler, "IPUMS USA: Version 14.0 [dataset]," 2023.

Sinha, Arnab, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-June Hsu, and Kuansan Wang, "An Overview of Microsoft Academic Service (MAS) and Applications," in "Proceedings of the 24th International Conference on World Wide Web" 2015, pp. 243–246.

Webb, Michael, "The Impact of Artificial Intelligence on the Labor Market," *Working Paper*, 2019.

Zhestkova, Yulia, "Technology Cluster Dynamics and Network Structure." PhD dissertation, The University of Chicago 2021.

APPENDIX

(For Online Publication)

A Details in Text Matching

This appendix reports the details of the matching between capital goods and patent data. Table A1 displays the share of tools found in Wikipedia for the NIPA categories.

Table A1: Share of Tools Found in Wikipedia

NIPA	Description	Found in Wikipedia (%)
20	Electrical transmission, distribution, and industrial apparatus	73.08%
4	Computers and peripheral equipment	69.64%
30	Furniture and fixtures	63.16%
27	Ships and boats	62.86%
40	Service industry machinery	60.00%
11	Office and accounting equipment	54.22%
29	Other equipment	53.13%
41	Electrical equipment, n.e.c.	53.10%
19	General industrial including materials handling equipment	52.03%
13	Fabricated metal products	49.62%
5	Communication equipment	48.59%
22,25	Trucks, buses, and truck trailers + autos	48.57%
14	Engines and turbines	46.81%
36	Construction machinery	44.44%
33	Agricultural machinery	42.86%
9	Nonmedical instruments	40.19%
10	Photocopy and related equipment	37.66%
18	Special industry machinery, n.e.c.	35.29%
39	Mining and oilfield machinery	31.13%
17	Metalworking machinery	30.61%
28	Railroad equipment	30.00%
6	Medical equipment and instruments	26.07%
26	Aircrafts	14.29%

Using the crosswalk between UNSPSC and NIPA from Caunedo et al. (2023), we assign a two-digit NIPA category code to each of the 4,180 tools. Then, we calculate the share of tools that are found in Wikipedia for each NIPA category. Table A1 shows that electronics, furniture, and machinery are more likely to be found in Wikipedia, while mining, medical equipment, and aircraft are less likely to be found in Wikipedia.

Table A2 shows the share of patents matched with at least one tool across time periods and patent classes. Patent classes are the one-digit codes of IPC. Patents on chemistry are rarely matched to tools, whereas patents on engineering have higher matching rates than others.

Table A2: Patent-Tool Matching Rates across Patent Class and Period

Patent Class	Matching Rate (%)			
	1970–1980	1980–1990	1990–2000	2000–2015
Human necessities	22.50	22.07	22.69	18.38
Transportation	33.19	32.75	33.48	26.80
Chemistry	6.74	7.31	8.18	9.06
Textile	28.53	30.61	31.36	26.19
Construction	32.14	31.30	31.87	24.29
Engineering	43.23	42.47	42.99	34.49
Physics	28.06	27.60	25.50	20.86
Electricity	27.48	27.74	25.87	21.87

Notes. This table presents the share of patents matched with at least one tool by period and patent class (IPC one-digit level).

A.1 Examples

Example 1 demonstrates how texts are matched. Texts are transformed into vectors of words, and the similarity score measures the distance between these vectors. The similarity score is higher if the two texts share more bigrams, pairs of consecutive words.

Example 1: Example of Text Matching between Patent and Wikipedia

Patent: Systems, apparatuses and methods for reading an amino acid sequence (10139417)
system apparatus method reading amino acid sequence embodiment present disclosure relate amino acid modified amino acid peptide protein identification sequencing mean example electronic detection individual amino acid small peptide

Wikipedia: Protein sequencer
protein sequencing practical process determining amino acid sequence part protein peptide may serve identify protein characterize post translational modification typically partial sequencing protein provides sufficient information one sequence tag identify reference database protein sequence derived conceptual translation gene

Notes. This figure shows an example of an abstract of the matched patent and Wikipedia article of the capital good. The blue text is the common words between two texts.

Example 2 shows a selected list of capital goods and their matched patents. The description of the capital goods is aligned with the titles of the matched patents.

Example 2: Example of Matched Capital Goods and Title of Patents

Capital Goods	Title of Patent
Battery chargers	Power tool, battery, charger and method of operating the same
Belt conveyors	Conveyor belt assembly
Cash registers	Theft proof cash drawer assembly
Desktop computers	Method and system for managing windows desktops in a heterogeneous server environment
Glass cutters	Discrete glass sheet cutting
Satellite phone	Communication system with direct link to satellite
Sewing machine needles	Multiple-needle sewing machine
Smoke detectors	Smoke detector system for a house
Tire pressure gauge	Tire pressure control system, tire pressure control device and tire pressure control method
Touch screen monitors	Technologies for interacting with computing devices using haptic manipulation

Notes. This table shows examples of matched capital goods and the title of patents.

B Measures of Capital Stock and User Costs

This appendix briefly explains how we impute the series of capital bundles and their user costs. Occupation-specific capital stock and user costs are calculated using procedures in Caunedo et al. (2023). Each occupation has a set of capital goods in the UNSPSC. These goods are converted to the NIPA capital category using the crosswalk in Table 1 of the Online Appendix for Caunedo et al. (2023). The 2012 fixed-price capital stock series is used to measure the quantity of capital. The capital intensity of an occupation o for the NIPA capital category n is first defined by the number of capital goods in the UNSPSC from "Tools Used" that are mapped into n. Let $\#\text{Capital}_o^{s,n}$ ($\#\text{Capital}_o^{d,n}$) denote the number of task-similar (task-dissimilar) capital goods in the UNSPSC and K_{int} the fixed-cost capital stock (based on the fixed price in 2012 USD) of industry i and NIPA capital category n in time t. Then, the capital stock of occupation o, industry i, and capital good n is imputed as

$$x_{siont} = \frac{l_{iot}\#\text{Capital}_o^{s,n}}{\sum_p l_{ipt}\#\text{Capital}_p^{s,n} + \sum_p l_{ipt}\#\text{Capital}_p^{d,n}} K_{int} \tag{A1}$$

$$x_{diont} = \frac{l_{iot}\#\text{Capital}_o^{d,n}}{\sum_p l_{ipt}\#\text{Capital}_p^{s,n} + \sum_p l_{ipt}\#\text{Capital}_p^{d,n}} K_{int}. \tag{A2}$$

l_{iot} is the number of occupations in industry i, occupation o, and time t. Thus, capital stocks are prorated between occupations with an intensity-weighted number of workers. If an occupation in an industry is missing from the O*NET and thus does not have any tool, the average intensity of tools in the industry is assigned to the occupation to adjust the capital stock. However, this occupation is not included in the regression analysis.

The price deflator is calculated as the ratio between current-cost and fixed-cost capital stock from the BEA and used as a capital price index. Depreciation rates are computed from depreciated capital stock data of the BEA. Specifically, the BEA depreciation rate d_{int} is calculated as the ratio of the depreciated capital stock in a year to the average between the capital stock evaluated at the end of the year and the capital stock evaluated at the end of the previous year. Because BEA-reported depreciation measures reflect both physical and economic depreciation, the physical depreciation rate is calculated using the following equation.

$$1 - \delta_{int} = (1 - d_{int}) \frac{q_{int}/\lambda_t^c}{q_{int-1}/\lambda_{t-1}^c} \tag{A3}$$

In this equation, λ_t^c is the price of consumption, and q_{int} is the price deflator. The user cost of capital category n for industry i and year t also comes from Caunedo et al. (2023) that follows Jorgenson (1963).

$$\lambda_{int}^k = \frac{q_{int}}{\lambda_{t-1}^c} \left[R - \left(1 - \bar{\delta}_{int}\right) \frac{q_{int}^k/\lambda_n^c}{q_{int-1}^k/\lambda_{t-1}^c} \right]. \tag{A4}$$

$R = 1.03$ is the gross return on a safe asset, and $\bar{\delta}_{int}$ is the average (physical) deflation rate of capital category n in industry i and the decade group t belongs to. If $t = 1980, \ldots, 1989$, $\bar{\delta} = \sum_{t=1980}^{1989} \delta_{int}$ with δ_{int} the annual deflation rate. We use $\lambda_t^c = 1$.

The quantity index of capital type $j = s, d$ for occupation o and industry i in year t is given as the following equation.

$$k_{jiot} = k_{jiot-1} e^{\kappa_{jiot}^k}, \quad k_{jio1980} = \sum_n x_{jion1980} \tag{A5}$$

$$\kappa_{jiot}^k = \sum_n \frac{\lambda_{int}^k x_{jiont}}{\sum_{n'} \lambda_{in't}^k x_{jion't}} \kappa_{int}^k. \tag{A6}$$

κ_{int}^k is the growth rate of capital category n. Thus, κ_{jiot}^k is the expenditure-weighted average growth rate of capital type j. Unlike Caunedo et al. (2023), we normalize the occupation-level stock, not the user costs of each occupation and industry, with the level in 1980. We take this approach because we are interested in the cross-sectional differences in capital stock and user cost series at the occupation level.

The user cost for the capital bundle is computed as follows.

$$r_{jiot} = \frac{\sum_n \lambda_{int}^k x_{jiont}}{k_{jiot}}. \tag{A7}$$

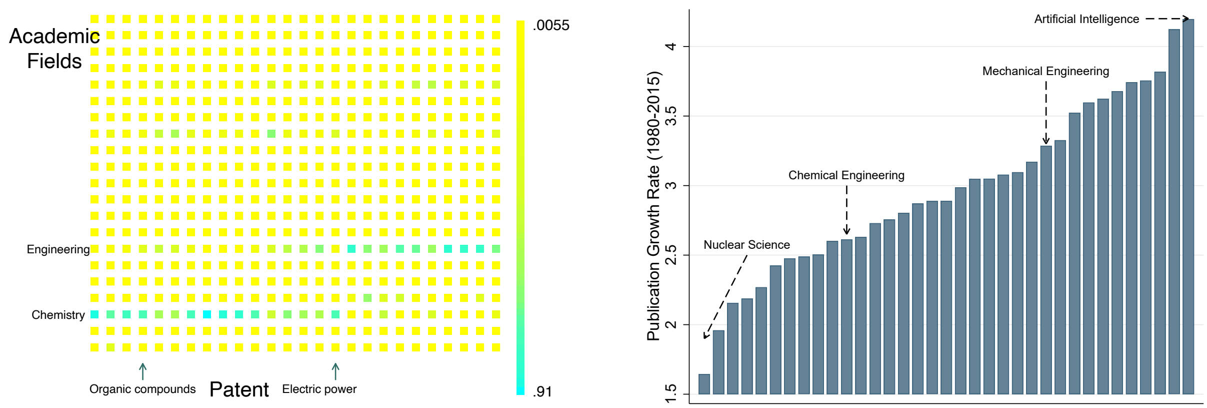
The equation uses the zero profit condition (1).

C Instrumental Variables

This appendix shows the variations that are used to formulate the shift-share instruments in the regression exercise and explains the equations that calculate the instruments.

C.1 Academic Publication Shock

Figure A1: Citation Share and Publication Growth Rate



A. Share of Citations B. Growth Rate of Publications

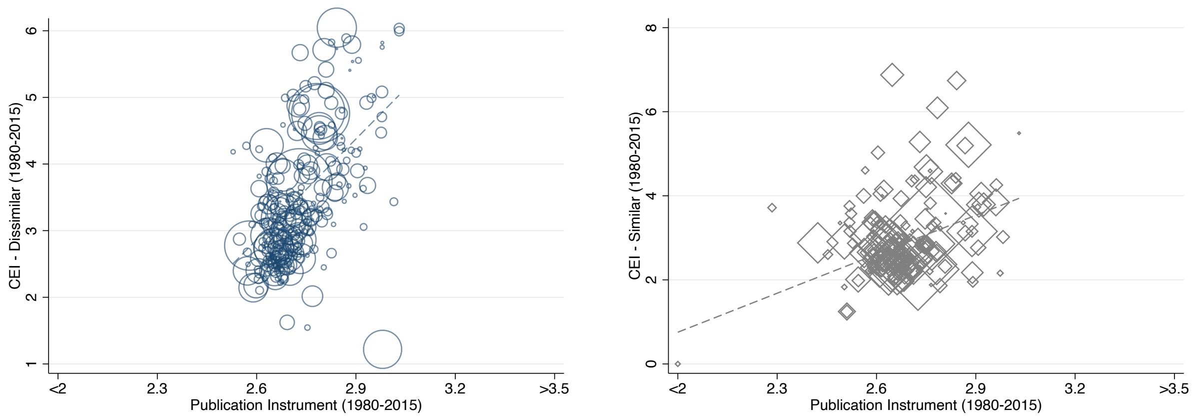
Notes. Panel A plots the share of citations from patent technology classes (row) to academic fields (column) in 1970–1980. The graph only contains the IPC classes that have more than 50,000 citations to science in the entire period. When the color gets closer to blue, it has a higher citation share. Panel B displays the growth rates of publications between 1980–2015 in different academic fields. Publication data comes from MAG and includes publications associated only with European institutions.

The left panel of Figure A1 plots υ_{pf}, showing the variation of citation share over patent classes. Engineering and chemistry are the fields that receive the most citations from patents. The right panel of Figure A1 displays the growth rates of publications in various academic fields. The fields with the highest growth rates include artificial intelligence, information systems, hardware, software engineering, and control systems.

$s^{Pat.}_{jiop}$ in Equation (17) is defined as below:

$$s^{Pat.}_{jiop} = \sum_{n} \left(\frac{\lambda^{k}_{jion} x_{jion} \#\text{Patent}_{jon}}{\sum_{n'} \lambda^{k}_{jion'} x_{jion'} \#\text{Patent}_{jon'}} \sum_{u \in \mathbb{U}(j,o,n)} \frac{1}{\#\text{Capital}^{j,n}_{o}} \left(\frac{\widehat{\#\text{Patent}_{up}}}{\sum_{p'} \widehat{\#\text{Patent}_{up'}}} \right) \right), \quad \text{(A8)}$$

Figure A2: CEI and Publication Instrument



A. CEI-d

B. CEI-s

Notes. Panel A plots CEI-d, and Panel B plots CEI-s over the publication instrument. Each circle corresponds to an occupation code in OCC1990, and the size of a circle corresponds to the stock value of capital in 1980. CEI is residualized with industry fixed effects.

where $\frac{\lambda_n x_n P_n}{\sum_{n'} \lambda_{n'} x_{n'} P_{n'}}$ is the expenditure-adjusted patent share of NIPA capital category n for capital bundle k_{jio}. $\#\,\text{Patent}_{jion}$ is the number of patents matched with the capital good n for jio. $\mathbb{U}(j, o, n)$ is the set of UNSPSC u of type j used by occupation o corresponding to n. $\#\text{Capital}_o^{j,n}$ is the number of capital goods in the UNSPSC that are classified as NIPA category n and is the cardinality of $\mathbb{U}(j, o, n)$. Lastly, $\#\,\widehat{\text{Patent}}_{up}$ is the number of patents in class p that are matched to the UNSPSC u. Thus, we give a higher weight to a patent class if more patents matched to the occupation have the patent class, the capital goods linked to the patent class are more representative in NIPA capital categories, and the NIPA category associated with the patent class accounts for more in expenditures and knowledge stock in the pre-period, 1980.

Then, the publication instrument is calculated as follows.

$$z_{jio}^{Pub.} = \sum_p s_{jiop}^{Pat.} \sum_f \upsilon_{pf} \Delta \log(\mathcal{P}_f) \,. \tag{A9}$$

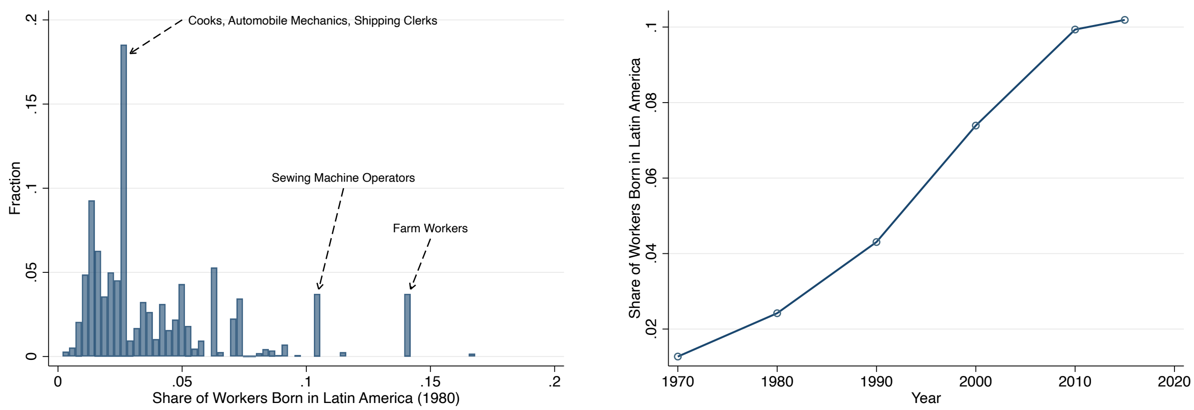
where $\mathcal{P}_f$ is the number of publications in field f.

Figure A2 displays the scatter plots between CEI measures and the resulting academic publication instruments, $z_{jio}^{Pub.}$, at the occupation level. The publication instruments are positively associated with the CEI measures, both for task-similar and task-dissimilar capital.

C.2 Immigration Shock

In order to construct an exogenous shifter in labor supply, we use trends in Latin American immigration and heterogeneous exposures to Latin American immigration. From 1980 to 2015, the population of Latin American-born workers in the US surged eightfold, while the number of US-born workers doubled. As a result, the share of workers born in Latin America in the total employment of the United States increased from 2.3% in 1980 to 10% in 2015, as shown in Panel A of Figure A3.

Figure A3: Share of Workers Born in Latin America



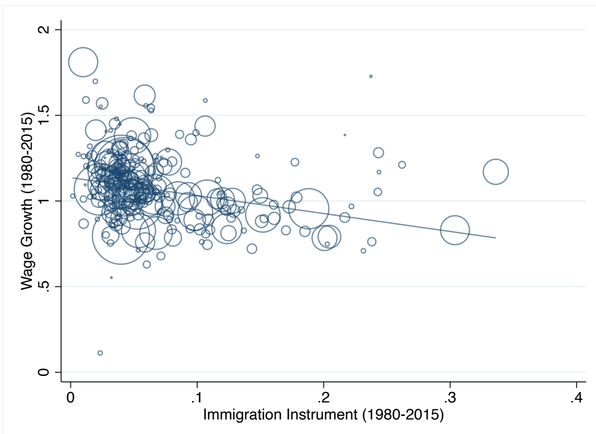
A. Share of Immigrants B. Growth of Immigrants

Notes. Panel A plots the share of workers born in Latin America in 1980 at the occupation level and draws the histogram of the observations. Each occupation is weighted by the number of workers in 1980. Panel A plots the share of workers in the US who were born in Latin America over time.

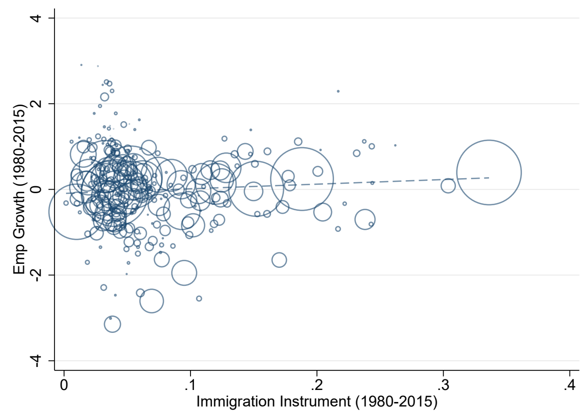
Immigrants from Latin America are likely to have comparative advantages different from those of US-born workers, influencing their choice of occupation differently. Panel B of Figure A3 shows the histogram of the share of workers from Latin America in 1980 for different occupations, with the weight of each occupation based on its employment numbers that year. The proportion of Latin American workers differs significantly between occupations. For example, in 1980, 13.5% of the farm workers were from Latin America, while less than 0.2% of the speech therapists were born in the region. Consequently, a surge in Latin American immigration would disproportionately affect the labor supply in certain occupations, such as farm workers.

Figure A4 plots the immigration instrument with the wage and employment growth rate at the occupation level. The immigration instrument is negatively correlated with wage growth while positively associated with employment growth. This suggests that our instrument serves as a shifter for labor supply.

Figure A4: Wage Change and Immigration Instrument



A. Wage Growth



B. Employment Growth

Notes. This figure plots the wage and employment growth rate of occupations between 1970 and 2015 with an immigration instrument. Each circle corresponds to an occupation code in OCC1990, and the size of a circle corresponds to the capital stock in 1980. Wage and employment growth are residualized with industry fixed effects.

C.3 Import Shock

Import shocks are constructed at the type(j)-industry(i)-occupation(o) level, as defined in Equation (19). The capital expenditure share in 1980 serves as an exposure measure, while the growth rate of import value is used as a shifter. The growth rate of import value is sourced from Comtrade data from 1980 to 2015. The Comtrade data records yearly bilateral trade flows at the product level. We use the imported data from all countries to the U.S. at the four-digit SITC Rev. 2 level. The SITC codes are manually matched with the NIPA capital categories. Table A3 displays the mapping.

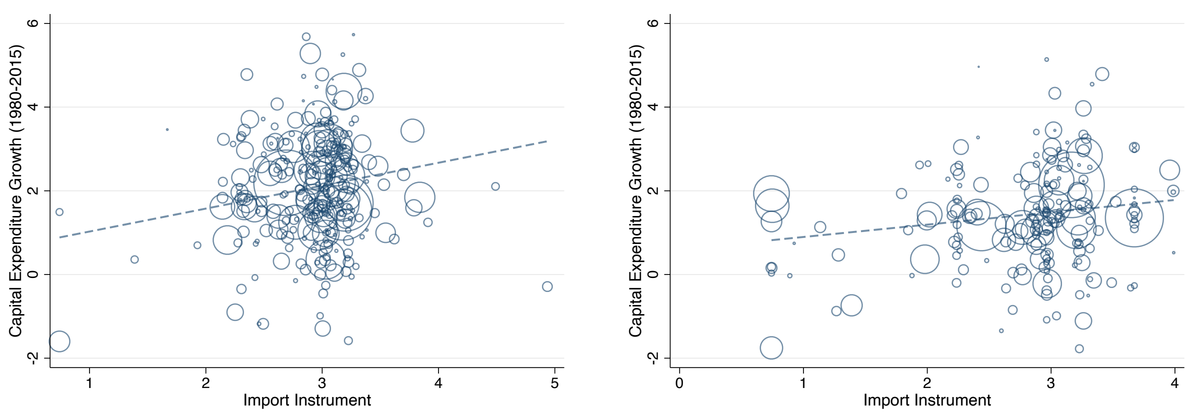
The left panel of Figure A5 displays the variation in the capital expenditure shares, and the right panel shows plots the distribution of growth rates of imported capital goods at the SITC Rev. 2 level. These figures suggest that the import instruments have sufficient variation both from the share and the shifter. Figure A6 plots the relationship between the import instruments and capital expenditure growth rates at the occupation level. Capital expenditure is calculated as the product of the stock and the user cost of capital goods. These results suggest that the import instrument is positively associated with capital expenditure, satisfying the relevance condition for the instrument.

Figure A5: Import Instrument and Capital Expenditure

A. Capital expenditure share in 1980 B. Growth rate of Imported value

Notes. Panel A shows the capital expenditure share in 1980 across two-digit occupation codes by capital type in the X-axis and NIPA capital category in the Y-axis. Capital expenditure is calculated as the stock value multiplied by the user cost of the capital. Panel B displays the distribution of growth rates for the import value of capital goods from 1980 to 2015. The data is sourced from Comtrade, where we aggregate four-digit SITC product codes into NIPA categories.

Figure A6: Import Instrument and Capital Expenditure



A. Task-dissimilar capital B. Task-similar capital

Notes. Panel A plots import instruments for task-dissimilar capital, and Panel B for task-similar capital over capital expenditure. Capital expenditure is stock values multiplied by user costs of capital. Each circle corresponds to an occupation code in OCC1990, and the size of a circle corresponds to the stock value of capital in 1980. The capital expenditure growth is residualized with industry fixed effects.

Table A3: Mapping Between NIPA Categories and HS Codes

NIPA code	Description of NIPA code	HS code	Description of HS code
4	Computers and peripheral equipment	8471	Automatic data processing machines and units thereof
4	Computers and peripheral equipment	8473	Parts for use with data processing machines
5	Communication equipment	8517	Line telephony and telegraphy apparatus, including telephones
5	Communication equipment	8525	Transmission apparatus for radio-telephony or broadcasting
5	Communication equipment	8526	Radar, radio navigational aid, and radio remote control apparatus
5	Communication equipment	8529	Parts for transmission and reception apparatus
6	Medical equipment and instruments	9018	Instruments for medical, surgical, dental, or veterinary use
6	Medical equipment and instruments	9019	Mechano-therapy, massage, and psychological aptitude-testing apparatus
6	Medical equipment and instruments	9021	Orthopedic appliances, including crutches, splints, and braces
6	Medical equipment and instruments	9022	X-ray apparatus and other medical diagnostic imaging apparatus
9	Nonmedical instruments	9031	Instruments for measuring or checking geometric quantities
9	Nonmedical instruments	9026	Instruments for measuring or checking liquid or gas flow, level, or pressure
9	Nonmedical instruments	9027	Instruments for physical or chemical analysis
9	Nonmedical instruments	9032	Automatic regulating or controlling instruments
10	Photocopy and related equipment	8443	Printing machinery and machines for uses ancillary to printing
10	Photocopy and related equipment	8442	Machinery and apparatus for printing plate preparation
11	Office and accounting equipment	8470	Calculating machines; accounting and ticketing machines
11	Office and accounting equipment	8472	Other office machines
13	Fabricated metal products	7308	Structures and parts of structures of iron or steel
13	Fabricated metal products	7326	Other articles of iron or steel
13	Fabricated metal products	7616	Other articles of aluminum
13	Fabricated metal products	7419	Other articles of copper
14	Engines and turbines	8411	Turbojets, turbopropellers, and other gas turbines
14	Engines and turbines	8406	Steam turbines and other vapor turbines
14	Engines and turbines	8407	Spark-ignition reciprocating or rotary internal combustion engines
14	Engines and turbines	8408	Compression-ignition internal combustion piston engines
17	Metalworking machinery	8456	Machine tools for working any material by removal of material
17	Metalworking machinery	8461	Machine tools for planing, shaping, slotting, and broaching
17	Metalworking machinery	8462	Machine tools for working metal by forging, hammering, or die-stamping
17	Metalworking machinery	8463	Other machine tools for working metals or carbides
18	General industrial equipment	8479	Machines and mechanical appliances with individual functions
18	General industrial equipment	8428	Lifting, handling, loading, or unloading machinery
18	General industrial equipment	8431	Parts suitable for use with lifting or moving machinery
20	Electrical equipment	8535	Electrical apparatus for switching electrical circuits
20	Electrical equipment	8536	Electrical apparatus for switching or protecting electrical circuits
20	Electrical equipment	8544	Insulated wire, cable, and other electric conductors
20	Electrical equipment	8504	Electrical transformers, static converters, and inductors
22	Trucks, buses, and truck trailers / Autos	8703	Motor cars and other motor vehicles for transporting people
22	Trucks, buses, and truck trailers / Autos	8704	Motor vehicles for the transport of goods
22	Trucks, buses, and truck trailers / Autos	8716	Trailers and semi-trailers; other vehicles
22	Trucks, buses, and truck trailers / Autos	8706	Chassis fitted with engines for motor vehicles
26	Aircraft	8802	Aircraft, including helicopters and airplanes
26	Aircraft	8803	Parts of aircraft or spacecraft
27	Ships and boats	8901	Vessels for transport of persons or goods
27	Ships and boats	8903	Yachts and other vessels for pleasure or sports
27	Ships and boats	8904	Tugs and pusher craft
28	Railroad equipment	8601	Rail locomotives powered by external sources
28	Railroad equipment	8602	Rail locomotives powered by an internal combustion engine
28	Railroad equipment	8607	Parts for railway or tramway locomotives
30	Furniture and fixtures	9401	Seats (except barber, dental, or similar chairs)
30	Furniture and fixtures	9403	Other furniture and parts thereof
33	Agricultural machinery	8432	Agricultural, horticultural, or forestry machinery
33	Agricultural machinery	8433	Harvesting, threshing, and other agricultural machines
33	Agricultural machinery	8436	Other agricultural, horticultural, forestry machinery
36	Construction machinery	8429	Self-propelled bulldozers, excavators, and road rollers
36	Construction machinery	8430	Other moving, grading, leveling, scraping, and boring machinery
39	Mining and oilfield machinery	8430	Other moving, grading, leveling, scraping, and boring machinery
40	Service industry machinery	8476	Automatic goods vending machines
40	Service industry machinery	8451	Machinery for laundering, drying, cleaning, or ironing textiles
88	Software	8523	Discs, tapes, and other recorded media
88	Software	8524	Recorded media for sound or video reproduction

Notes. This table displays the manual mapping between NIPA capital categories and HS codes (1992).

D Robustness

D.1 Alternative Thresholds and Citation Measures

The baseline threshold is set at the 90th percentile of the occupation-capital similarity score distribution to distinguish task-similar and task-dissimilar capital. This threshold gives statistically significant differences to the CEI-s and CEI-d measures in the structural regression. Table A4 compares the estimation results across different thresholds for task-similar capital. The estimates of structural parameters are robust overall to different thresholds and the citation-based measures of CEI, except when the threshold is set at the 89th percentile. When the threshold is set at the 89th percentile and the set of task-similar capital is extended, the elasticity of substitution of task-similar capital is now lower than the elasticity of substitution between occupational services.

Table A4: Estimation Results with Different Thresholds and Citations

	ρ_s	ρ_d	σ	γ_p	γ_s	γ_d
90th Percentile	3.02	1.55	2.27	-0.469	-0.59	0.86
	(0.97)	(0.24)	(0.24)	(0.058)	(0.18)	(0.31)
89th Percentile	1.41	0.63	2.38	-0.32	-1.31	2.02
	(0.19)	(0.27)	(0.33)	(0.12)	(0.48)	(0.53)
91st Percentile	2.54	1.50	2.38	-0.33	-0.52	1.12
	(0.46)	(0.13)	(0.18)	(0.11)	(0.10)	(0.13)
Citation	4.73	2.11	2.72	-0.34	-0.54	1.15
	(1.74)	(0.41)	(0.16)	(0.12)	(0.10)	(0.12)

Notes. This table shows estimates of structural parameters in Table 5 with alternative CEI measures and publication instruments. ρ_s (ρ_d) is the elasticity of substitution between task-similar (task-dissimilar) capital and labor. σ is the elasticity of substitution between different occupational services. γ_p is the coefficient of CEI on user costs of capital. γ_s (γ_d) is the coefficient of CEI-s (-d) on occupational service demand shifter. The rows labeled the 90th, 89th, and 91st Percentiles mean the percentiles of task-capital similarity scores used to calculate thresholds for similar-dissimilar distinction. The baseline uses the 90th percentile. The row labeled Citation uses the number of citations on the patents to measure CEI.

D.2 Alternative Nesting

In the paper, we first combine labor with task-similar capital in the inner CES composite and then task-dissimilar capital in the outer composite. We now change the order of nesting and report the results. The results are quantitatively similar to the baseline, except for ρ_s.

Table A5: Parameter Estimates Under Alternative Nesting

	ρ_s	ρ_d	σ	γ_s	γ_d
Estimate	5.44	1.73	2.44	-0.70	1.36
SE	(3.36)	(0.19)	(0.18)	(0.09)	(0.10)

Notes. This table shows the estimates and standard errors of the regression equations (15) and (16). ρ_s (ρ_d) is the elasticity of substitution between task-similar (task-dissimilar) capital and labor. σ is the elasticity of substitution between different occupational services. γ_s (γ_d) is the coefficient of CEI-s (-d) on occupational service demand shifter.

D.3 Reweighting Using Wikipedia Missing Rates

Since some tools are not found in Wikipedia and are excluded from our calculation, CEI can be underestimated if certain occupations use capital goods with higher missing rates more intensively. To address this issue, we give weights to capital goods by the inverse of the finding rate in Wikipedia in the Appendix Table A2. Table A6 shows the estimation results.

Table A6: Parameter Estimates Under Reweighting

	ρ_s	ρ_d	σ	γ_p	γ_s	γ_d
Estimate	1.242	1.060	1.145	-0.338	-0.549	0.642
SE	(0.506)	(0.122)	(0.265)	(0.121)	(0.142)	(0.175)

Notes. This table shows the estimates and standard errors of the regression equations (15) and (16) when the tools are weighted by the inverse of the finding rate in Wikipedia. ρ_s (ρ_d) is the elasticity of substitution between task-similar (task-dissimilar) capital and labor. σ is the elasticity of substitution between different occupational services. γ_s (γ_d) is the coefficient of CEI-s (-d) on the demand shifter for occupational services.

E Industry Variations of CEI

We report the results of the variance decomposition of CEI measures between occupations and industries in Table A7. The variation of CEI-s predominantly comes from occupation-level heterogeneity in the lists of capital goods. However, a non-negligible fraction of variation comes from industry-level heterogeneity in capital composition within each occupation for CEI-d. Due to industry-level heterogeneity in occupational composition, between-industry variations of CEI measures are also sizeable in the second panel.

Table A7: Variance Decomposition

	Similar		Dissimilar	
	CEI	IV	CEI	IV
Between Occupation	0.923	0.961	0.786	0.882
Within Occupation	0.077	0.039	0.214	0.118
#Occupations	211	-	291	-
Between Industry	0.421	0.606	0.373	0.467
Within Industry	0.579	0.394	0.627	0.533
#Industries	59	-	59	-

Notes. This table shows the variations of the CEI measures and their instruments between occupations and industries.

F Equations for Counterfactual Equilibrium

The counterfactual exercise aims to derive the counterfactual equilibrium without CEI in 1980-2015. We only allow # Patent$_{jion}$ to be at their levels in 1980 and let other terms in demand and supply equations stay at their original levels in 2015. The total employment **L** is also fixed at its level in 2015.

Following Caunedo et al. (2023), we assume a CES aggregator to make capital bundles with $\phi = 1.13$ as the elasticity of substitution.

$$k_{jio} = \left(\sum_n \xi_{jion}^{\frac{1}{\phi}} x_{jion}^{\frac{\phi-1}{\phi}} \right)^{\frac{\phi}{\phi-1}} \tag{A10}$$

Additionally, two following equations are needed to run the counterfactual equilibrium.

$$1 = \frac{\alpha_i}{\alpha_j} \left(\frac{\mu_{io}}{\mu_{jo}} \right)^{\frac{1}{\sigma}} \left(\frac{Y_i}{Y_j} \right)^{\frac{1}{\sigma}-1} \left(\frac{y_{io}}{y_{jo}} \right)^{\frac{1}{\rho_d}-\frac{1}{\sigma}} \left(\frac{\tilde{\Theta}_{io}}{\tilde{\Theta}_{jo}} \right)^{\frac{\rho_d-\rho_s}{\rho_s\rho_d}} \left(\frac{l_{io}}{l_{jo}} \right)^{-\frac{1}{\rho_d}} \tag{A11}$$

$$Y_i = l_{io} \left(\sum_o \mu_{io}^{\frac{1}{\sigma}} \left(\frac{l_{io}}{l_{i0}} \right)^{\frac{\sigma-1}{\sigma}} \tilde{y}_{io}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} = l_{i0}\tilde{Y}_i \tag{A12}$$

Equation (A11) is derived from the first order conditions of cost minimization with respect to l_{io} and l_{jo}, respectively. Equation (A12) expresses industrial outputs as a linear function of l_{io}, labor input of a reference occupation 0, and $\tilde{Y}_i$ that only depends on the ratio of labor inputs relative to a reference occupation 0. The manager (OCC1990 = 22) is used as the reference occupation.

By combining Equations (A11) and (A12), the following equation is derived.

$$1 = \frac{\alpha_i}{\alpha_j} \left(\frac{\mu_{io}}{\mu_{jo}} \right)^{\frac{1}{\sigma}} \left(\frac{\tilde{Y}_i}{\tilde{Y}_j} \right)^{\frac{1}{\sigma}-1} \left(\frac{\tilde{y}_{io}}{\tilde{y}_{jo}} \right)^{\frac{1}{\rho_d}-\frac{1}{\sigma}} \left(\frac{\tilde{\Theta}_{io}}{\tilde{\Theta}_{jo}} \right)^{\frac{\rho_d-\rho_s}{(\rho_s-1)\rho_d}} \left(\frac{l_{io}}{l_{jo}} \right)^{-1} \tag{A13}$$

We use this equation to pin down the industry-level employment of an occupation.

G CEI and User Costs of Capital

In this section, we explore in more detail the relationship between CEI and user costs of capital and the role of user cost changes in labor market trends. First, we summarize changes in user costs associated with CEI. Then, we quantify how user cost changes alone contribute to heterogeneous labor market trends for different occupations using the structural model estimated in Section 4.

Table A8 summarizes the actual and counterfactual user costs of capital for selected NIPA asset types in 1980, 1990, 2000, and 2015. CEI variables are redefined using patents granted until 1990, 2000, and 2015 relative to patents granted until 1980. Computers experience the sharpest decline in user costs over time. The baseline estimate $\gamma_p = -0.41$ is used to calculate counterfactual user costs. Without CEI, computer user costs still show the largest decline, but the magnitude is much more muted. In fact, along with communication and photocopy equipment, computer user costs are the most heavily affected by CEI. Office equipment, metal products, and automobiles have the lowest CEI measures, and thus their user costs are the least affected by the absence of CEI.

Table A8: CEI and User Costs of Capital

	Actual				Without CEI			
	1980	1990	2000	2015	1980	1990	2000	2015
Computers	6.90	3.24	0.71	0.51	6.90	5.15	2.06	1.53
Communication eq.	0.46	0.47	0.34	0.18	0.46	0.64	0.69	0.45
Medical eq.	0.15	0.19	0.20	0.34	0.15	0.26	0.33	0.60
Photocopy eq.	0.29	0.32	0.24	0.38	0.29	0.50	0.57	1.13
Office eq.	0.33	0.38	0.38	0.54	0.33	0.42	0.51	0.72
Metal products	0.08	0.09	0.10	0.32	0.08	0.11	0.14	0.44
Engines	0.07	0.07	0.10	0.29	0.07	0.08	0.15	0.54
Aircraft	0.04	0.06	0.08	0.33	0.04	0.07	0.13	0.67
Electrical eq.	0.15	0.19	0.20	0.40	0.15	0.24	0.32	0.76
Autos	0.13	0.19	0.22	0.42	0.13	0.22	0.31	0.61

Notes. This table shows the average and counterfactual user costs of capital for selected categories of NIPA capital over time. Counterfactual average user costs are calculated using patent measures from 1980 relative to 2015 with a price elasticity of $\gamma_P = -0.41$. 'eq.' means equipment.

Table A9 summarizes the changes in user costs for similar and dissimilar capital over occupation groups distinguished by task scores, labor share, and wage level. After price-level adjustments, all occupation groups have declines in user costs of capital. Abstract and high-wage occupations experience larger reductions in user costs of similar and dissimilar capital. Routine task scores are not monotonically associated with changes in user costs, and occupations with high labor shares have greater declines in user costs of dissimilar capital but smaller reductions in similar capital.

Table A9: User Cost Changes

Similar			Dissimilar		
Low	Middle	High	Low	Middle	High
Panel A. Across Abstract Score					
-0.29	-0.23	-0.37	-0.34	-0.53	-0.54
Panel B. Across Routine Score					
-0.13	-0.34	-0.20	-0.49	-0.48	-0.54
Panel C. Across Labor Share in 1980					
-0.41	-0.25	-0.17	-0.38	-0.48	-0.64
Panel D. Across Wage in 1980					
-0.24	-0.24	-0.36	-0.36	-0.51	-0.57

Notes. This table summarizes the changes in user costs of similar and dissimilar capital used by occupations in the first (low), second to fourth (mid) and fifth (5Q) quintiles in the distributions of task scores, labor share, and wage level in 1980. Changes in user costs adjust for the change in the CPI between 1980 and 2015.

We repeat the counterfactual exercise with the level of capital user costs fixed at their levels in 1980. Table A10 compares the resulting task bias of changes in the counterfactual equilibrium. The effect of user costs on task-biased labor market changes is small because the estimates for the elasticities of substitution are close to each other and their values are close to one. In addition, the difference in changes in user costs between occupation groups is small.

Table A11 summarizes the changes in aggregate labor share and their decomposition in the data and in counterfactual equilibria without changes in user costs. User costs alone can generate the historical decline in labor share, larger than the reduction in the CEI exercise.

Table A10: Task-Biased Labor Market Changes without Changes in User Costs

	Abstract		Routine	
	Employment	Wage	Employment	Wage
Actual Change	0.17	0.44	-0.26	-0.33
Without CEI (η=0.3)	0.17	0.43	-0.26	-0.34
Without CEI (η=1)	0.17	0.43	-0.26	-0.34

Notes. This table shows the actual and counterfactual regression coefficients of the wage and employment growth rates at the occupation level in the occupational task scores of Autor and Dorn (2013) as in Equation (20). The counterfactual equilibrium fixes the user costs of all capital inputs at their levels in 1980. The columns $\eta = 0.3$ and $\eta = 1$ set the elasticity of the occupational labor supply at 0.3 and 1, respectively.

Table A11: Declines in Labor Share without User Cost Changes

	Within	Between	All
Actual Change	-8.64	-1.24	-9.87
Without CEI (η=0.3)	-0.73	0.84	0.11
Without CEI (η=1)	-0.70	0.85	0.15

Notes. This table shows the actual and counterfactual changes in the labor share between 1980 and 2015. The counterfactual equilibrium fixes the user costs of all capital inputs at their levels in 1980. Changes in labor share within and between occupations are derived in Equation (21). The rows labeled 'Without CEI ($\eta = 0.3$)' and 'Without CEI ($\eta = 1$)' set the elasticity of occupational labor supply at 0.3 and 1, respectively.

Lastly, Table A12 summarizes labor market polarization in a counterfactual equilibrium with changes in user costs only. The impact of user cost changes is small on polarization statistics.

Table A12: Polarization without User Cost Changes

	Employment			Wage		
	Low	Middle	High	Low	Middle	High
Actual Change	0.17	-0.58	0.41	-0.03	-0.54	0.57
Without CEI (η=0.3)	0.17	-0.58	0.41	-0.06	-0.50	0.56
Without CEI (η=1)	0.16	-0.58	0.42	-0.05	-0.52	0.57

Notes. This table shows the actual and counterfactual growth rates of the wage and employment growth of occupations grouped by their wages in 1980. The counterfactual equilibrium fixes the user costs of all capital inputs at their levels in 1980. The wage and employment changes are subtracted from the mean and divided by the standard deviation of the changes at the occupation level in each case. Columns labeled Low and High denote occupations with 1980 wage levels in the first and fifth quintiles, respectively. The columns labeled middle denote occupations between the two quintiles. The rows labeled 'Without CEI ($\eta = 0.3$)' and 'Without CEI ($\eta = 1$)' set the elasticity of occupational labor supply at 0.3 and 1, respectively.

H Computers and Robots

Computers and robots have been considered one of the most important technological changes in the labor market (Autor et al., 2003; Acemoglu and Restrepo, 2018; Burstein et al., 2019). This section summarizes the importance of computers and robots in CEI measures. Then, we repeat the counterfactual exercise after setting only innovations unrelated to computers fixed at the level in 1980. Since robots account for a negligible fraction of CEI, we do not repeat the counterfactual exercise that sets only innovations related to robots fixed at the level in 1980.

Table A13: Share of Computer and Robot in Matched Patents (%)

	1970	1980	1990	2000	2010
Computer	1.56	3.05	9.86	10.13	8.92
Robot	0.27	0.31	0.22	0.22	0.21

Notes. This table displays the shares of computers and robots in patents matched to capital goods in the UNSPSC. The numbers are in percent. The column titles indicate the decades for which the patents were applied, except under 2010. Column 2010 collects patents applied from 2011 to 2015.

Table A13 shows the share of computers and robots in patents matched to capital goods in the UNSPSC. A capital good is considered a computer if the commodity title is in the 'computer equipment and accessories' family in the UNSPSC (43210000). A capital good is considered a robot if the title has the words "automatic," "robot," or "drone." Computers accounted for 1.6% of patents in the 1970s, but their importance steadily increased over time. About 10 percent of patents applied after the 1990s are computer-related. On the other hand, robots make up only 0.2-0.3% of patents in all columns. Due to the scarcity of robot-related patents, we exclude robots in the following exercises.

Next, Table A14 summarizes the share of computers in similar and dissimilar capital stock across occupation groups. Over time, the share of computer stock increased primarily in dissimilar capital, and computers accounted for almost 30% of capital stock in 2015. However, the increase was nearly uniform across task scores. Occupations with low labor shares still have low shares of computers in total capital in 2015, and occupations with high labor shares tend to have higher computer intensities at all times.

Table A14: Computer Stock Intensity across Occupation Groups

	Similar			Dissimilar		
	Low	Middle	High	Low	Middle	High
	Panel A. Across Abstract Score					
1980	0.00	3.62	0.34	1.73	6.50	6.14
2015	2.73	4.55	2.72	23.11	27.92	22.86
	Panel B. Across Routine Score					
1980	0.00	3.76	0.06	2.35	6.47	5.66
2015	0.03	5.21	3.76	25.61	24.99	29.36
	Panel C. Across Labor Share in 1980					
1980	0.29	3.28	1.10	0.03	3.17	17.66
2015	1.79	3.66	6.36	0.67	28.03	45.17
	Panel D. Across Wage in 1980					
1980	0.00	3.64	0.44	1.50	7.28	4.03
2015	2.75	4.87	1.85	14.56	31.89	20.12

Notes. This table summarizes the share of computers in similar and dissimilar capital over occupations grouped by task scores, labor shares, and wage levels. The columns labeled Low and High represent the occupations in the first and fifth quintiles, respectively, while those labeled Middle cover occupations within the second to fourth quintiles. Numbers in percent.

Table A15 demonstrates the importance of computers in CEI for the occupation groups in Table A14. Computers made up a small fraction of patents on dissimilar capital for all occupation groups in 1980. In 2015, because computers are knowledge-intensive capital goods, computers represented a large fraction of patents related to dissimilar capital, except for occupations with low labor shares and low wages in the first quintile. Still, since computers are used by all occupations, their importance in CEI is relatively stable across occupations grouped by task scores.

To quantify the impact of computer-related innovation on labor market changes, we re-calculate the counterfactual equilibrium without changes in computer-related patents between 1980 and 2015. All other patents remain in 2015. Table A16 displays the coefficient estimates of task scores on wage and employment changes in counterfactual equilibrium. Without increases in computer-related patents over time, labor market changes would have been less biased towards abstract and against routine occupations, but the

Table A15: Shares of Computer-Based Patents across Occupation Groups

	Similar			Dissimilar		
	Low	Middle	High	Low	Middle	High
	Panel A. Across Abstract Score					
1980	0.00	5.31	1.80	1.62	8.73	17.30
2015	3.14	6.49	2.65	10.60	27.70	27.09
	Panel B. Across Routine Score					
1980	0.00	4.24	0.48	11.92	9.90	2.19
2015	0.00	5.10	7.42	24.22	22.42	28.63
	Panel C. Across Labor Share in 1980					
1980	0.18	4.62	2.39	0.65	9.29	15.98
2015	0.75	5.31	10.87	2.31	25.93	40.63
	Panel D. Across Wage in 1980					
1980	0.00	4.75	1.99	3.15	8.89	14.68
2015	3.30	6.37	2.39	8.44	28.34	27.29

Notes. This table summarizes the share of computers in matched patents of occupations grouped by task scores, labor shares, and wage levels. The columns labeled Low and High represent the occupations in the first and fifth quintiles, respectively, while those labeled Middle cover occupations within the second to fourth quintiles. The numbers are in percent.

magnitude of the change is much smaller than in Table 6. Thus, the contribution of computer-related innovations is limited. This is again because almost all occupations use computers, and computers are almost uniformly important in CEI measures.

Table A16: Task-Biased Labor Market Changes without Computer-based CEI

	Abstract		Routine	
	Employment	Wage	Employment	Wage
Actual Change	0.17	0.44	-0.26	-0.33
Without CEI (η=0.3)	0.18	0.22	-0.25	-0.14
Without CEI (η=1)	0.16	0.25	-0.22	-0.17

Notes. This table shows the actual and counterfactual regression coefficients of the wage and employment growth rates at the occupation level in the occupational task scores of Autor and Dorn (2013) as in Equation (20). The counterfactual equilibrium uses CEI measures that fix only patents unrelated to computers at their levels in 1980. The rows labeled 'Without CEI ($\eta = 0.3$) and 'Without CEI ($\eta = 1$) set the elasticity of occupational labor supply at 0.3 and 1, respectively.

Table A17 shows the changes in the labor shares without computer-based CEI. Similarly to the task exercise, computer-based CEI generates a much smaller fraction of actual declines in labor shares.

Table A17: Declines in Labor Share without Computer-based CEI

	Within	Between	All
Actual Change	-8.64	-1.24	-9.87
Without CEI (η=0.3)	-7.18	-0.80	-7.97
Without CEI (η=1)	-7.21	-0.76	-8.34

Notes. This table shows the actual and counterfactual regression coefficients of the wage and employment growth rates at the occupation level in the occupational task scores of Autor and Dorn (2013) as in Equation (20). The counterfactual equilibrium uses CEI measures that fix only patents unrelated to computers at their levels in 1980. The rows labeled 'Without CEI ($\eta = 0.3$) and 'Without CEI ($\eta = 1$) set the elasticity of occupational labor supply at 0.3 and 1, respectively.

Lastly, Table A18 describes counterfactual changes in employment and wage growth between occupations grouped by their wages in 1980. The results imply that computers played a small role in generating labor market polarization. According to Table A18, when computer-based innovations are back to the level of 1980, the employment growth of high-wage occupations in the fifth quintile barely changes. This again comes from computers being widely used as dissimilar capital and the intensity of computers in CEI being nearly uniform over the second to the fifth quintiles of 1980 wage levels.

Table A18: Polarization without Computer-based CEI

	Employment			Wage		
	Low	Middle	High	Low	Middle	High
Actual Change	0.20	-0.61	0.41	-0.04	-0.53	0.57
Without CEI (η=0.3)	0.21	-0.61	0.40	0.01	-0.36	0.35
Without CEI (η=1)	0.21	-0.59	0.38	-0.00	-0.39	0.39

Notes. This table shows the actual and counterfactual growth rates of the wage and employment growth of occupations grouped by their wages in 1980. The counterfactual equilibrium uses CEI measures that fix only patents unrelated to computers at their levels in 1980. The wage and employment changes are subtracted from the mean and divided by the standard deviation of the changes at the occupation level in each case. Columns labeled Low and High denote occupations with 1980 wage levels in the first and fifth quintiles, respectively. The columns labeled middle denote occupations between the two quintiles. The rows labeled 'Without CEI ($\eta = 0.3$)' and 'Without CEI ($\eta = 1$)' set the elasticity of occupational labor supply at 0.3 and 1, respectively.