Explaining Job Polarization: The Role of Heterogeneity in Capital Intensity*

WAN-JUNG CHENG†

We propose a new perspective that job polarization stems from the interaction of the decrease in the relative price of capital goods, the heterogeneity in capital intensity of job task production, and the complementarity of job tasks in final goods production. First, we construct a measure of occupation-level capital intensities and document that the tasks of middle-skill workers tend to be more capital intensive. Second, we build a task-based model with two goods sectors and three job tasks, where the job task production differs in capital intensity and how a worker's skill is utilized. The model demonstrates that when there is technological progress in investment technology, employment shifts away from capital-intensive tasks, and the relative wages are driven down, implying that a decreasing price of capital goods predicts job polarization. A quantitative analysis suggests that the model can account for approximately one-half of the employment polarization and approximately one-half of the upper tail of the wage polarization in the U.S. between 1980 and 2010.

JEL Classification: E24, E25, J21, J31, O33

Keywords: Job polarization, Task-based model, Capital intensity, Technological change

*This paper is based on the first chapter of my Ph.D. dissertation at Washington University in St. Louis. I am grateful to my dissertation committee members Costas Azariadis, Limor Golan, Yongseok Shin, Ping Wang, and David Wiczer for their advice on this project. I also thank comments from Georg Graetz, Jang-Ting Guo, Charles Ka Yui Leung, and audiences at the 70th European Meeting of the Econometric Society, 2017 North American Summer Meeting of the Econometric Society, 2017 Asian Meeting of the Econometric Society, 2016 Fall Midwest Macroeconomics Meeting, Academia Sinica, and National Taiwan University. Errors are my own.

†Correspondence: Institute of Economics, Academia Sinica, 128 Academia Road, Section 2, Nangang, Taipei 11529, Taiwan. Email: wjcheng@econ.sinica.edu.tw.

1 INTRODUCTION

Job polarization, which consists of employment and wage polarization, is regarded as a pervasive phenomenon. Over the past few decades, employment shares have shifted away from middle-skill occupations toward high- and low-skill occupations, which is known as employment polarization. Meanwhile, the average wages of middle-skill occupations have been growing noticeably more slowly than those of high- and low-skill occupations, which is referred to as wage polarization. Job polarization has been widely documented in advanced economies, including the United States and Western European countries. For example, Figure 1 shows this pattern in the U.S. by plotting the change in the employment share and the percentage change in the relative wage for each occupation between 1980 and 2010. Both graphs are U-shaped: high-skill and low-skill occupations experienced growth in both employment shares and relative wages, while middle-skill occupations experienced the opposite change.[1]

In this paper, we propose a new perspective to explain the job polarization. We first present a new finding that tasks performed by middle-skill occupations are more capital intensive.[2] Based on this observation, we propose a new mechanism through which job polarization can be induced by a well-documented shock: technological progress in the production of investment goods.[3] Because occupations are heterogeneous in capital intensities, they respond differently to this economy-wide investment shock. When different occupations are complementary, a positive shock to the investment technology induces a shift in employment away from the capital-intensive occupation along with a decrease in its relative wage—namely, employment polarization and wage polarization.

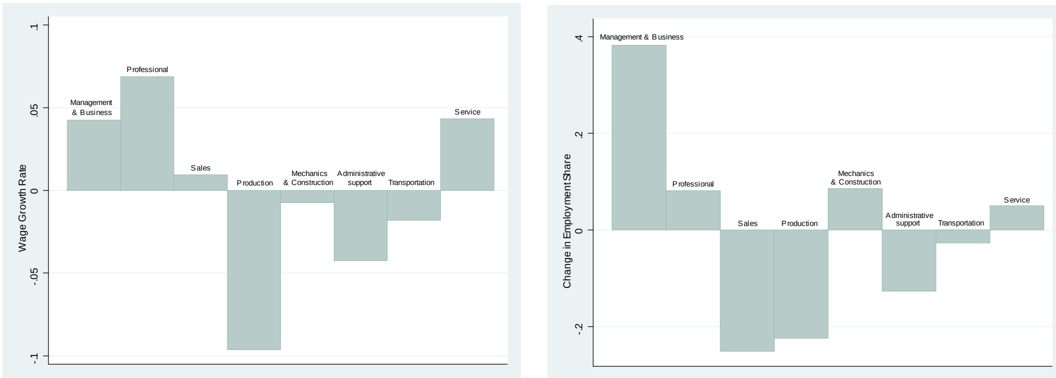
In the literature, the mainstream hypothesis is that job polarization results from a routine-biased technological change (RBTC), which refers to the process of comput-

[1]These two labor market trends have attracted researchers' attention because they cannot be successfully explained by the canonical skill premium model, which has been used to explain between-group inequality. Therefore, there is a need for a new theory that can explain the trends in employment and earnings across the distribution defined with finer skill groups. See Acemoglu and Autor (2011) for detailed discussions.

[2]Hereinafter, we use the terms "tasks" and "occupations" interchangeably.

[3]The literature (Fisher [2006]; Justiniano et al. [2009]; Karabarbounis and Neiman [2014]) has identified a trend break in the price of investment goods relative to consumption goods—suggesting technological progress in the production of investment goods—circa 1980, when job polarization is believed to appear.

Figure 1: Employment and Wage Polarization



Left panel: the change in the employment share. Right panel: the percentage change in the relative wage. Occupations are ranked on the x-axis by their skill level, from lowest to highest, where the skill rank is proxied by the average wage of workers in that occupation in 1980. The employment share is defined as the ratio of the number of workers in the occupation to the total number of workers. The relative wage is defined as the ratio of the average wage in the occupation to the average wage of all workers.

erization of routine tasks. Because routine tasks are mainly performed by middle-skill occupations, in the literature, RBTC is captured by an occupation-specific technological shock that raises the productivity of middle-skill workers, and thus, RBTC shifts labor out of middle-skill occupations.[4] The literature argues that job polarization derives from a shock that affects only particular types of occupations, whereas we argue that it is an economy-wide investment technology shock that plays this role.

We construct a measure of occupation-level capital intensity by utilizing the measures of industry-level capital intensities that have been available in the literature. We document that not only are the tasks of middle-skill occupations more capital intensive than those of low- and high-skill occupations but also that such capital intensity gaps widen over time. We build a task-based model that consists of three job tasks and two final goods sectors.[5] Workers and capital are used to produce job task outputs, which are in turn used to produce final consumption and investment goods. The three job tasks, corresponding to high-skill, middle-skill, and low-skill occupations, are produced using capital and workers' skills with Cobb-Douglas technology. Job task production differs in capital intensity and how a worker's skill

[4]The RBTC hypothesis, also referred to as the routinization hypothesis, was formulated by Autor et al. (2003). See Autor (2013) for detailed reviews.

[5]The framework was introduced by Acemoglu and Autor (2011).

is utilized; the latter results in positive sorting between workers' skill levels and task production technologies. That is, workers with the highest skill levels sort themselves into the tasks with the highest return on skills, and workers with lowest skill levels sort themselves into the tasks with the highest return on raw labor. The consumption and investment goods are produced by combining the three types of tasks with CES technology, and as in the literature, job tasks are assumed to be complementary in producing final goods.

In the presence of technological progress in producing investment goods, the price of investment goods declines, and aggregate capital thereby increases. The higher capital intensity of the middle-skill task production indicates that its technology employs capital in a more productive manner. Thus, an equiproportional increase in capital input would raise the output of the middle-skill task to a greater extent. Because different tasks are complements, it must be the case that a smaller fraction of aggregate capital is allocated to the middle-skill task than before; otherwise, excessive middle-skill task outputs would be produced. Accordingly, this feature induces labor to flow out of the middle-skill task, implying employment polarization. The model also features a decrease in the average wage of the capital-intensive task relative to the average wages of the other tasks, implying wage polarization. Alongside job polarization, the labor productivity of middle-skill workers grows more rapidly than that of high- and low-skill workers. In contrast to the literature, the relative increase in the productivity of middle-skill workers in our model is an equilibrium result coexisting with job polarization. Moreover, the widened capital intensity gaps further amplify the abovementioned effects.

In the quantitative analysis, we assess the extent to which the model can explain the employment and wage polarization in the U.S. from 1980 to 2010. In other words, we quantify how much of the changes in employment shares and relative wages can be explained by the decline in the relative price of investment goods given the heterogeneity in capital intensity of task production. We construct a measure of occupation-level capital intensities and use the estimates for calibration. The year 1980 is chosen as the base year, as the early 1980s are considered the starting point of polarization. We simulate two equilibria, one for the base year 1980 and the other for 2010. All parameters are assumed to be time-invariant except for the capital intensities and two other parameters that govern the relative price of the investment good and the rental rate of capital. Thus, the second equilibrium represents the model's predictions of employment and wages in 2010.

Our calibrated model can well capture the job polarization in the U.S. experienced in recent decades. The model explains one-half of the decrease in the middle-skill employment share, one-third of the increase in the high-skill employment share, and 90% of the increase in the low-skill employment share. Moreover, the model also explains one-half of the upper tail of wage polarization. Out of the total explained fraction, the decomposition analysis indicates that approximately 40% is attributable to the decrease in the price of capital goods, and the change in capital intensities accounts for the other 60%. Furthermore, middle-skill workers' labor productivity tripled, which is approximately 1.5 times the labor productivity growth in high- and low-skill occupations. We also compare our results with the RBTC channel through a counterfactual analysis, where we shut down our channel and, instead, let the productivity of middle-skill workers grow exogenously. We show that for the RBTC channel to match the extent to which our channel can explain the decrease in middle-skill employment, the productivity of middle-skill workers would need to grow nearly 3 times as high as the productivity growth of high- and low-skill workers. Moreover, the RBTC channel always generates an increasing trend for the aggregate labor share, while our channel induces a decreasing trend, which is consistent with the data.

RELATED LITERATURE

Job polarization has been observed in many developed countries, and the phenomenon is believed to have existed since at least the early 1980s.[6] The routinization (or RBTC) hypothesis was originally formulated by Autor et al. (2003), who present evidence that computers substitute for workers in routine tasks but complement workers in nonroutine tasks. Theoretical studies model routinization in various ways. In Acemoglu and Autor (2011) and Autor and Dorn (2013), routinization results from technological progress in the production of computer capital, and tasks performed by middle-skill workers are relatively substitutable by computer capital. In Goos et al. (2014), routinization is modeled as a trend in production costs whereby routine task production enjoys decreasing costs of its non-labor inputs, but their paper does not model workers' skills and hence does not address the issue of wage polarization. Although the modeling strategies are distinct, these papers share

[6]For the United States, see Autor et al. (2006), Autor et al. (2008), Acemoglu and Autor (2011), and Autor and Dorn (2013). For the United Kingdom, see Goos and Manning (2007). For other Western European countries, see Goos et al. (2009) and Goos et al. (2014).

a common feature: they study how a shock that directly impacts a certain set of occupations has effects on the whole economy. Similarly, Stokey (2016) theoretically studies the general equilibrium effects of technological changes within a specific set of occupations in an environment where the technology complements workers' human capital. While most studies focus on the task content of occupations, Bárány and Siegel (2017) emphasize a sectoral explanation: workers with specific skill levels are tied to specific sectors. These authors argue that job polarization results from structural change.

Our paper is also related to studies on the interplay of investment-specific technological progress and long-run labor market trends. Karabarbounis and Neiman (2014) show that the significant decline in the labor share observed across countries since 1980 can be well explained by the decrease in the relative price of investment goods, which reflects investment-specific technological progress. Krusell et al. (2000) study the evolution of the skill premium and demonstrate that the decrease in the price of investment can well account for the skill premium in the U.S., which has increased significantly since 1980. Note that 1980 not only marks the trend break in the decline in investment prices, but several salient labor market phenomena also manifested themselves in this period—for example, the decline in the labor share, the increase in the skill premium, and job polarization. To the best of our knowledge, we are the first to explain job polarization by a shock that affects the economy as a whole (i.e., an investment shock) through a channel whereby occupations are affected to different extents owing to their heterogeneity (in how intensively they use capital). One implication that we want to capture is that job polarization is not a unique phenomenon but is closely related to other labor market trends that have been believed to result from progress in investment technology.

The heterogeneous capital intensity setting in our paper is related to that of Acemoglu and Guerrieri (2008), in which capital deepening induces a reallocation of capital and labor away from the capital-intensive sector. Our paper features a similar reallocation of capital and labor across occupations in the presence of capital deepening: an increase in aggregate capital stock decreases the shares of capital and labor in more capital-intensive occupations. By considering heterogeneous skills, our paper can further address how the movements of relative wages are related to capital intensities. The mechanism in our model also has similar features to those in Ngai and Pissarides (2007), where a low elasticity of substitution across sectors leads to shifts of employment from sectors with high productivity growth to sectors with low productivity growth. In our model, the investment technology shock leads to high productivity growth for middle-skill tasks relative to others. As a result, employment shifts from middle-skill tasks to low- and high-skill tasks, which have lower productivity growth.

2 Model

We consider a task-based model in which there are two hierarchies in the production structure: producing final goods requires combining different types of job tasks, and job tasks are produced using human capital (skill) and physical capital as inputs. The economy is populated by a continuum of heterogeneous workers. Each worker is endowed with one unit of time and skill $h \in \mathbb{R}^+$, which is drawn from the distribution $F(h)$.

2.1 Final Goods

There are three types of job tasks, $j \in J = \{1, 2, 3\}$, and two final goods sectors, the consumption good C and investment good X. Final goods are produced by combining three types of job tasks according to a CES technology, where $Y_{c,j}$ and $Y_{x,j}$ denote the quantity of job task j used in the production of the consumption good and investment good, respectively:

$$C = \left[\sum_{j \in J} \lambda_j^{\frac{1}{\eta}} Y_{c,j}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}, \tag{1}$$

$$X = \xi_x \left[\sum_{j \in J} \lambda_j^{\frac{1}{\eta}} Y_{x,j}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}. \tag{2}$$

As in Karabarbounis and Neiman (2014), $\xi_x \in \left[\underline{\xi}, \overline{\xi}\right]$ denotes the technology level in the production of the investment good relative to the consumption good.[7] The task intensity $\lambda_j \in (0, 1)$ is assumed to be the same across sectors with $\sum_{j \in J} \lambda_j = 1$. The parameter $\eta \in \mathbb{R}^+$ is the elasticity of substitution between tasks in the production

[7] Our setting is also theoretically similar to that in Grossman et al. (2017), in which investment-specific technological progress is a component of capital-augmenting technological progress.

of final goods. When $\eta < 1$, tasks are complementary in producing final goods; this complementarity property is widely adopted in the literature on job polarization and will also be the working assumption in this paper.

ASSUMPTION 1 *Tasks are complementary in the production of final goods:*

$$0 < \eta < 1.$$

Let q_j denote the price of task j and p_c and p_x denote the prices of the consumption and investment good, respectively. Given p_c, p_x, and q_j, final goods producers solve

$$\max_{Y_{c,j}} \; p_c C - \sum_{j \in J} q_j Y_{c,j}, \tag{3}$$

$$\max_{Y_{x,j}} \; p_x X - \sum_{j \in J} q_j Y_{x,j}. \tag{4}$$

2.2 Tasks

Job tasks are produced when one producer and one worker are paired, using the worker's skill h and physical capital k as inputs. The job task production technology is Cobb-Douglas and takes the following form:

$$y_j(h) = \tilde{e}_j(h)^{1-\alpha_j} k^{\alpha_j}. \tag{5}$$

The capital intensity α_j is heterogeneous across tasks, and we make the following assumption in accordance with the empirical findings reported in Section 5.

ASSUMPTION 2 *The production of task* 2 *is more capital intensive:*

$$\alpha_2 > \alpha_1 \;\; \textit{and} \;\; \alpha_2 > \alpha_3.$$

This assumption reflects our observations that tasks performed by middle-skill occupations are more capital intensive. More importantly, this feature serves as the key heterogeneity between tasks and is the essential factor in explaining how an economy-wide shock would affect tasks of different magnitudes. Departing from the literature, we propose a novel view that job polarization results from a shock that affects the whole economy, rather than a shock that specifically hits certain tasks. The fact that tasks are affected differently, i.e., polarization, owes to the heterogeneity in how intensively the task relies on capital for production.

Workers contribute to task production through $\tilde{e}_j(h)$, the efficiency units of labor derived from a worker with skill h in the production of task j. Notice that $\tilde{e}_j(h)$ is a function of both j and h, indicating that how a worker's skill is utilized depends on which task (occupation) this worker performs. An intuitive specification of $\tilde{e}_j(h)$ is that the labor that a worker provides includes two components: human capital and physical labor, which we henceforth call skill and raw labor, respectively. Motivated by Herrendorf and Schoellman (2017), we assume a linear form:

$$\tilde{e}_j(h) = \gamma_j h + \theta_j \tag{6}$$

where $\gamma_j, \theta_j \in \mathbb{R}^+$. That is, the efficiency units of labor are derived linearly from the worker's skill and raw labor, with different rates of return depending on which task the worker performs. The parameter γ_j, referred to as the return on skills, captures the extent to which the production of task j utilizes the worker's skill. The parameter θ_j, referred to as the return on raw labor, captures the contribution per worker, regardless of skill level, to the production of task j. Both γ_j and θ_j are allowed to differ across tasks, indicating that task production differs in how the technology employs the two labor force characteristics of the workers. Specifically, we make the following assumption.

Assumption 3 *The return on skill (γ_j) is increasing in j, and the return on raw labor (θ_j) is decreasing in j.*

This assumption implies that high-h workers have a comparative advantage in high-j tasks and that low-h workers have a comparative advantage in low-j tasks. As will be shown in Proposition 1, Assumption 3 helps ensure positive sorting between skill h and task j in equilibrium: high-h workers endogenously work in high-j tasks, and vice versa. To facilitate the proof, we let γ_j and θ_j obey the following relationships:

$$\begin{aligned} \gamma_2 &= \left(1+\varepsilon_1^{\gamma}\right)\gamma_1 \quad ; \quad \gamma_3 = \left(1+\varepsilon_2^{\gamma}\right)\gamma_2 \\ \theta_1 &= \left(1+\varepsilon_1^{\theta}\right)\theta_2 \quad ; \quad \theta_2 = \left(1+\varepsilon_2^{\theta}\right)\theta_3 \end{aligned}$$

where ε_1^{γ}, ε_2^{γ}, ε_1^{θ}, $\varepsilon_2^{\theta} > 0$.

Let q_j denote the price of task j and $w_j(h)$ denote the wage paid to a worker with skill level h employed in task j. Given task price q_j, wage profile $w_j(h)$, and capital rental rate R, producers of task j solve the following problem:

$$\max_{h,k} q_j y_j(h) - w_j(h) - Rk. \tag{7}$$

Finally, let $\mathbb{O}_j$ denote the endogenous set of workers in task j. The total output of each task j, denoted by Y_j, can be allocated to the production of both the consumption good and investment good:

$$Y_{c,j} + Y_{x,j} = Y_j = \int_{h \in \mathbb{O}_j} y_j(h)\, dF(h); \tag{8}$$

Two points are worth noting here: first, the functional form of the wage profile $w_j(h)$ can potentially differ across tasks; second, the policy function for capital will be a function of h and should be denoted by $k_j(h)$ in equilibrium, as it can also potentially differ across tasks.

2.3 Workers

Workers are endowed with skill h and are heterogeneous in skill levels. Given the wage profile $w_j(h)$, each worker inelastically supplies one unit of time and chooses an occupation $\hat{\jmath} \in J$. Workers purchase the consumption good at price p_c and obtain utility from consumption. They also purchase the investment good at price p_x, use the investment good to augment their capital stock, and rent out capital at rental rate R. Workers can hold asset b that pays a real interest rate r and is in zero net supply.

Specifically, given an occupation j, a worker with skill h solves the following maximization problem:

$$V_j(h) = \max_{\{c_t, x_t, b_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t(h)),$$

subject to the budget constraint and the law of motion for capital

$$p_c c_t + p_x x_t + b_{t+1} - (1+r) b_t = w_j(h) + R k_t,$$

$$k_{t+1} = (1-\delta) k_t + x_t.$$

Knowing $V_j(h)$, the worker with skill h chooses the occupation whereby he obtains the highest utility:

$$\hat{\jmath}(h) = \arg\max_{j \in J} V_j(h).$$

3 EQUILIBRIUM

3.1 Definition of Equilibrium

The steady-state equilibrium of this economy consists of the prices of final goods $\{p_c, p_x\}$, the prices of tasks $\{q_j\}_{j\in J}$, the capital rental rate R, the wage profile $w_j(h)$, the quantities of final goods $\{C, X\}$, the quantities of tasks $\{Y_j\}_{j\in J}$, the allocations $\{Y_{c,j}, Y_{x,j}\}_{j\in J}$, and the occupation choice schedule $\hat{\jmath}(h)$ such that:

1. (Maximization) Producers of final goods and job tasks solve the maximization problem (3), (4), (7), respectively. Workers solve the occupation choice and maximization problem defined in Subsection 2.3.
2. (Markets clear)
   (a) The labor market clears: $\mathbb{O}_j = \{h \mid \hat{\jmath}(h) = j\}$.
   (b) The job task markets clear: for all $j \in J$, (8) holds.
   (c) The goods market clears:

$$C + p_x X = \sum_{j\in J} \int_{\mathbb{O}_j} w_j(h)\, dF(h) + RK^D \tag{9}$$

   where $K^D = \sum_{j\in J} \int_{\mathbb{O}_j} k_j(h)\, dF(h)$ and $k_j(h) = \arg\max_k \, q_j y_j(h) - w_j(h) - Rk$.
   (d) The capital market clears: $K^S = K^D$, where $K^S = \frac{1}{\delta}X$.

The key equilibrium properties are presented in the remainder of this section, and detailed derivations are presented in Appendix A.

3.2 Price of Final Goods

Let p_c and p_x denote the prices of the consumption good and investment good, respectively. The consumption good is considered the numeraire, and hence, its price is normalized to unity. We can obtain the solution that, in the competitive market, the price equals the marginal cost of production:

$$p_c = \left[\sum_{j\in J} \lambda_j {q_j}^{1-\eta}\right]^{\frac{1}{1-\eta}} = 1.$$

Similarly, the price of the investment good also equals the marginal cost of production:

$$p_x = \frac{1}{\xi_x}\left[\sum_{j\in J} \lambda_j {q_j}^{1-\eta}\right]^{\frac{1}{1-\eta}} = \frac{1}{\xi_x}. \tag{10}$$

In equilibrium, the relative price of the investment good, defined as $\frac{p_x}{p_c}$, is equal to $\frac{1}{\xi_x}$, the inverse of the technology in the investment good sector. In steady state, the rental rate of capital is negatively correlated with ξ_x:

$$R = \frac{1}{\xi_x}\left(r+\delta\right). \tag{11}$$

When there is an increase in ξ_x, referred to as a technological improvement in the investment good sector relative to the consumption good sector, there is a decline in both the relative price of the investment good and the rental rate of capital. The capital rental rate would also decrease if there is a decrease in the real interest rate r.

3.3 Wage Profile

By solving the task producer's problem, we can show that the equilibrium wage profile is a linear function of $\tilde{e}_j\left(h\right)$:[8]

$$w_j(h) = \varpi_j\left(q_j, \xi_x\right) \times \tilde{e}_j\left(h\right). \tag{12}$$

where $\tilde{e}_j\left(h\right)$ is defined in (6) and $\varpi_j\left(q_j, \xi_x\right) = {q_j}^{\frac{1}{1-\alpha_j}} {\xi_x}^{\frac{\alpha_j}{1-\alpha_j}} \left(1-\alpha_j\right)\left(\frac{\alpha_j}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}}$. Alternatively, $w_j(h)$ can be expressed as $w_j(h) = a_j + b_j h$ such that

$$\begin{aligned} a_j &= \theta_j \varpi_j\left(q_j, \xi_x\right), \\ b_j &= \gamma_j \varpi_j\left(q_j, \xi_x\right). \end{aligned}$$

[8] See Appendix A for detailed derivations.

The linear form of the wage profile immediately implies that if $a_1 > a_2 > a_3$ and $b_3 > b_2 > b_1$ hold, there will be positive sorting between skill h and task j—namely, high-h workers endogenously sort into high-j tasks, and low-h workers endogenously sort into low-j tasks. Positive sorting requires that high-h workers have sufficient comparative advantages in high-j tasks (measured by the return on skills), and low-h workers have sufficient comparative advantages in low-j tasks (measured by the return on raw labor). Our analysis is summarized in the following proposition.

PROPOSITION 1 (POSITIVE SORTING) *Denote $\boldsymbol{\varepsilon} = \left(\varepsilon_1^{\gamma}, \varepsilon_2^{\gamma}, \varepsilon_1^{\theta}, \varepsilon_2^{\theta}\right)$. Under Assumption 3, there is an $\underline{\varepsilon}$ such that if $\boldsymbol{\varepsilon} > \underline{\varepsilon}$ we can find the thresholds $\hat{h}_1$ and $\hat{h}_2$ such that*

$$
\begin{aligned}
\hat{\jmath}(h) &= 1 \quad \text{if} \quad h < \hat{h}_1, \\
\hat{\jmath}(h) &= 2 \quad \text{if} \quad \hat{h}_1 < h < \hat{h}_2, \\
\hat{\jmath}(h) &= 3 \quad \text{if} \quad h > \hat{h}_2.
\end{aligned}
$$

Proof. See Appendix B.1. ■

This proposition immediately implies that we can express $\mathbb{O}_j$ as follows:

$$
\begin{aligned}
\mathbb{O}_1 &= \left\{h \mid h < \hat{h}_1\right\}, \\
\mathbb{O}_2 &= \left\{h \mid \hat{h}_1 < h < \hat{h}_2\right\}, \\
\mathbb{O}_3 &= \left\{h \mid h > \hat{h}_2\right\}.
\end{aligned}
$$

In equilibrium, workers endogenously sort into tasks according to their skill levels and the extent to which the task production technology utilizes workers' labor force components. Positive sorting means that high-skill workers will work in occupations with high returns on skills, while low-skill workers will work in occupations with high returns on raw labor, and those in between will be employed in middle-skill occupations. It also means that workers with the highest skills receive the highest wages, or more specifically, workers' skill levels are positively correlated with their wage levels. Note that this property thus rationalizes the use of wage levels as the proxy for skill levels in the empirical analysis in Section 5.

Given (5) and the $k_j(h)$ solved in Appendix A.2, we can also express the wage profile $w_j(h)$ as

$$
w_j(h) = q_j \frac{\partial y_j}{\partial \tilde{e}_j(h)} \tilde{e}_j(h).
$$

The wage paid for a skill-h worker in task j can be decomposed into three parts: the level of the efficiency units of skill $\tilde{e}_j(h)$, the marginal product of the efficiency units of skill $\frac{\partial y_j}{\partial \tilde{e}_j(h)}$, and the marginal revenue of the task output q_j. That is, the equilibrium wage is equal to the efficiency units of the worker's skill multiplied by the marginal contribution of that skill to the task producer's total revenue. At skill thresholds $(\hat{h}_1, \hat{h}_2)$, the equilibrium wages must equalize the contribution of the threshold worker's skill to the task producer's revenues between tasks.

$$w_1\left(\hat{h}_1\right) = w_2\left(\hat{h}_1\right);\;\; w_2\left(\hat{h}_2\right) = w_3\left(\hat{h}_2\right).$$

From the workers' perspective, the equilibrium wages must make the threshold workers indifferent across tasks. That is, a worker with skill $\hat{h}_1$ must be indifferent to being employed in either tasks 1 or 2, and similarly, a worker with skill $\hat{h}_2$ must be indifferent to being employed in either tasks 2 or 3.

4 COMPARATIVE STATICS

The aim of the comparative statics in this section is to analyze the impact of an increase in ξ_x, the technology level in the production of the investment good. Specifically, we focus on the changes in the employment shares and relative wages of the low-, middle-, and high-skill occupations. Positive sorting implies that $j \in \{1,2,3\}$ can thus denote the low-, middle-, and high-skill tasks, respectively. Let ES_j denote the share of employment in task j. ES_j is defined as

$$ES_j \equiv \int_{\mathbb{O}_j} dF(h).$$

According to Proposition 1, we can thus express the employment shares of low-skill, middle-skill, and high-skill occupations as follows:

$$\begin{aligned} ES_1 &= \int_0^{\hat{h}_1} dF(h), \\ ES_2 &= \int_{\hat{h}_1}^{\hat{h}_2} dF(h), \\ ES_3 &= \int_{\hat{h}_2}^{\infty} dF(h). \end{aligned} \tag{13}$$

Moreover, to facilitate the analysis in this section, we henceforth assume that the distribution of the skill follows $h \sim \text{log-normal}(\mu, \sigma)$.

4.1 Employment Polarization

We first examine the impact of increasing ξ_x on employment shares, and the analysis builds on the following proposition. This proposition states that the model predicts employment polarization if there is technological progress in the production of the investment good. In other words, the proposition also implies a close connection between job polarization and the decline in the price of the investment good, which we observe in the data. Assume that the conditions specified in Proposition 1 hold (i.e., $\varepsilon > \underline{\varepsilon}$), and we have the following proposition.

PROPOSITION 2 (EMPLOYMENT POLARIZATION) *Under Assumptions 1, 2, and 3, the technological progress in the production of the investment good (an increase in ξ_x) results in a decrease in the employment share of middle-skill occupations (ES_2) and increases in the employment shares of both low- and high-skill occupations (ES_1 and ES_3).*

Proof. See Appendix B.2. ■

The positive shock to the investment technology leads to an increase in the aggregate capital stock. This causes the output of the more capital-intensive task (middle-skill task) to increase relative to the output of less capital-intensive tasks (low- and high-skill tasks). In other words, an increase in ξ_x increases the productivity of all tasks but does so to the greatest extent for the capital-intensive task. Therefore, a low elasticity of substitution (η less than 1) across tasks leads to shifts of employment and capital away from the capital-intensive task, which experiences the fastest growth in productivity.

To better explain the mechanism, we derive the relationship among the relative prices of tasks:[9]

$$\frac{q_j}{q_{j'}} = \left(\frac{\lambda_j}{\lambda_{j'}}\right)^{\frac{1}{\eta}} \left(\frac{Y_j}{Y_{j'}}\right)^{-\frac{1}{\eta}} \quad \forall j, j' \in J. \tag{14}$$

Equation (14) states that the relative price of tasks is negatively related to the relative quantity of task outputs. When the elasticity of substitution η is less than one, a relative increase in task output will decrease that task's relative price more than propor-

[9] See Appendix A for detailed derivations.

tionately. This equation is helpful for understanding how employment polarization is induced. If aggregate capital increases and the shares of capital allocated to tasks remain constant, then the output of the capital-intensive task would grow by more than that of the other tasks because an equiproportional increase in capital would raise the output of the capital-intensive task to a greater extent. When $\eta < 1$, the increase in the output of the capital-intensive task causes its price to fall more than proportionately, inducing a reallocation of capital away from the capital-intensive task to the other tasks. Along with the reallocation of capital, labor also moves away from the capital-intensive task to the other tasks, implying employment polarization. We formally summarize the mechanism in the following Corollary.

COROLLARY 1 *Under Assumptions 1, 2, and 3, the technological progress in the production of the investment good (an increase in ξ_x) results in*

1. *Decreases in the relative prices of the middle-skill task ($\frac{q_2}{q_1}$ and $\frac{q_2}{q_3}$);*
2. *Increases in the relative quantities of the middle-skill task outputs ($\frac{Y_2}{Y_1}$ and $\frac{Y_2}{Y_3}$);*
3. *Decreases in the relative quantities of capital in the middle-skill task production ($\frac{K_2}{K_1}$ and $\frac{K_2}{K_3}$).*

Proof. See Appendix B.3 ■

4.2 Wage Polarization

Let $\overline{w}_j$ denote the average wage of the workers in occupation j. That is,

$$\overline{w}_j \equiv \frac{\int_{\mathbb{O}_j} w_j(h)\, dF(h)}{ES_j}. \tag{15}$$

The following proposition states that the technological progress in the production of the investment good can also induce decreases in the average wage of middle-skill workers relative to those of low-skill and high-skill workers.

PROPOSITION 3 (WAGE POLARIZATION) *Under Assumptions 1, 2, and 3, the technological progress in the production of the investment good (an increase in ξ_x) results in decreases in the average wage of middle-skill workers relative to those of both low-skill and high-skill workers ($\frac{\overline{w}_2}{\overline{w}_1}$ and $\frac{\overline{w}_2}{\overline{w}_3}$).*

Proof. See Appendix B.4. ■

Proposition 2 also implies that employment polarization is associated with an increase in $\hat{h}_1$ and a decrease in $\hat{h}_2$, that is, the inward movement of both skill thresholds. Specifically, the proof shows that

$$\frac{\partial\left(\frac{\overline{w}_2}{\overline{w}_1}\right)}{\partial \hat{h}_1} < 0 \text{ and } \frac{\partial\left(\frac{\overline{w}_2}{\overline{w}_3}\right)}{\partial \hat{h}_2} > 0.$$

That is, in the presence of employment polarization, wages also polarize. According to equation (42), the task-specific average wage $\overline{w}_j$ depends on both the task price q_j and the price of the investment good $\frac{1}{\xi_x}$, and how the relative wages would change thus depends on how the task prices and the investment good price change in a relative sense. On the one hand, for task 2, the capital-intensive task, increasing ξ_x raises its relative wage because an equiproportional increase in capital would increase the productivity of capital-intensive task production by more. On the other hand, when tasks are complementary, $\eta < 1$, the price of the capital-intensive task would fall more than proportionally in response to the increase in its productivity, as shown in 14. As a result, due to the complementarity, the fact that q_2 decreases more than proportionally leads not only to the outflow of labor from task 2 but also to the decrease in its relative wage, where the latter is because the effect of decreasing q_2 on $\overline{w}_2$ dominates the effect of the increasing ξ_x.

4.3 Task-specific Productivity

It has been argued in the literature that job polarization is induced by a task-specific technological shock that raises the productivity of middle-skill workers. This paper departs from the literature and proposes that job polarization is induced by an economy-wide technological shock that affects all tasks, but to different extents. The differences result from how intensively tasks rely on capital for production. The heterogeneity in capital intensity in task production not only causes job polarization but also generates differences in the change in task-level productivity. Specifically, we argue that the productivity of middle-skill workers indeed increases, but, in contrast to the literature, the increase is an endogenous result accompanied by job polarization. This subsection will demonstrate how task-specific labor productivity is related to task-specific capital intensity and affected by the shock to the investment good production technology.

The total output of task j, according to (8), is derived as follows:

$$Y_j = q_j^{\frac{\alpha_j}{1-\alpha_j}} \xi_x^{\frac{\alpha_j}{1-\alpha_j}} (r+\delta)^{-\frac{\alpha_j}{1-\alpha_j}} \alpha_j^{\frac{\alpha_j}{1-\alpha_j}} \left(\gamma_j \exp\left(\mu + \frac{1}{2}\sigma^2 \right) + \theta_j \right) ES_j,$$

where ES_j is the employment share of task j. The task-specific labor productivity, denoted by $\hat{y}_j$, is defined as

$$\hat{y}_j \equiv \frac{Y_j}{ES_j}.$$

The following proposition states that the technological progress in the production of the investment good also leads to increases in the productivity of middle-skill workers relative to those of low- and high-skill workers.

PROPOSITION 4 (TASK-SPECIFIC PRODUCTIVITY) *Under Assumptions 1, 2, and 3, the technological progress in the production of the investment good (an increase in ξ_x) results in increases in the labor productivity of middle-skill workers relative to those of both low-skill and high-skill workers ($\frac{\hat{y}_2}{\hat{y}_1}$ and $\frac{\hat{y}_2}{\hat{y}_3}$).*

Proof. See Appendix B.5. ■

Thus, the model predicts that the labor productivity of middle-skill occupations continues to grow along with the decrease in the investment good price. In contrast to the literature, where the task-specific shock that raises the productivity of middle-skill workers is the driving force of job polarization, our model implies that the increasing productivity of middle-skill workers is an equilibrium outcome. Specifically, both job polarization and the increase in the productivity of middle-skill tasks are driven by an economy-wide shock, given the observed heterogeneity in capital intensity across tasks.

Note that when ξ_x increases, the increase in the relative productivity of the middle-skill task (increases in $\frac{\hat{y}_2}{\hat{y}_1}$ and $\frac{\hat{y}_2}{\hat{y}_3}$) is in line with a decrease in the relative price of middle-skill task output (decreases in $\frac{q_2}{q_1}$ and $\frac{q_2}{q_3}$). The capital-intensive feature of middle-skill task production means that the technology can more productively use capital. Thus, an increase in the aggregate capital in the economy, induced by the advance in the investment technology, raises the productivity of middle-skill workers more. Since different tasks are complementary, the output prices of less capital-intensive tasks must increase to attract production factors. The outflow of capital and labor from the capital-intensive task explains employment polarization. Moreover, the decreases in $\frac{q_2}{q_1}$ and $\frac{q_2}{q_3}$ that induce the factor flow are the driving forces of wage polarization, as the decreases in prices are more than proportionate to the increases in output.

5 DATA AND MEASUREMENT

This section will show that the tasks held by middle-skill occupations tend to be more capital intensive. First, we categorize occupations into three skill groups based on one-digit occupation codes in the U.S. census and the classifications in the literature. Then, we construct a measure of occupation-level capital intensities and use them to estimate task-level capital intensities. Finally, we use empirical evidence to demonstrate how job polarization is related to capital intensities.

5.1 Classification of Skill Groups

First, occupations are categorized into nine major groups based on one-digit occupation codes and the availability of the industry-occupation matrix, which is needed to construct occupation-level capital intensity, in the census data. Then, occupations are ranked by their skill levels, which are approximated by occupational mean wages in 1980.[10] Finally, the nine occupations are further categorized into the three skill groups: high-, middle- and low-skill groups. The classification is shown in Table 1, which is consistent with the classification in the literature. The reported values in the wage column are the ratios of occupational mean wages to the mean wage of workers in 1980. The wage ratios in 2010 are also reported; the classification of the skill groups remains intact, although the ranking of occupations within the middle-skill group changes.

The high-skill group consists of management, business, financial, and professional occupations. The middle-skill group consists of production-related occupations, administrative support, and sales occupations. The low-skill group consists of physically demanding service occupations such as food service, cleaning, personal care, and protective service workers. It is clear that high-skill occupations are more human-capital intensive, and low-skill occupations are more physical-labor inten-

[10] Specifically, the wage data mentioned in this paper refer to the data on earnings.

Table 1: Occupation Groups

Skill group	Occupations	Wage (1980)	Wage (2010)
High skill	Management, business, financial	1.5435	2.1332
	Professional	1.3423	1.4519
Middle skill	Sales	1.1070	0.8284
	Production, operators, assemblers	0.9691	0.7604
	Installation, maintenance, mechanics	0.9107	0.9640
	Construction and extraction	0.8939	0.9878
	Office and administrative support	0.8631	0.7537
	Transportation and material moving	0.7667	0.7534
Low Skill	Service	0.5666	0.5946

Data source: U.S. census.

sive. The next subsection demonstrates that the middle-skill occupations are more capital intensive.

5.2 Capital Intensity Index

The first empirical contribution of this paper is to show that middle-skill occupations are more capital intensive than both high-skill and low-skill occupations. Specifically, an empirical estimate of α_j will be developed, and we will demonstrate that α_2 is larger than both α_1 and α_3. Although it is well known that the data on the capital share can be used to calibrate the capital intensities of industries, there are no direct data on the capital intensities of occupations. Therefore, an implied capital intensity needs to be constructed to form the estimate of α_j.

We first construct the capital intensity index for each occupation c, denoted by $\hat{\alpha}_c$. Then, the occupation-level capital intensities $\hat{\alpha}_c$ are used to form the estimates of task-level capital intensity α_j, which will be a weighted sum of $\hat{\alpha}_c$ over the occupations belonging to that task group. Let κ_d denote the measure of capital intensity of industry d and $\tau_{(c,d)}$ the employment share of occupation c workers who are employed in industry d. For each occupation c, the capital intensity is measured as follows:

$$\hat{\alpha}_c = \sum_d \tau_{(c,d)} \times \kappa_d.$$

The concept of $\hat{\alpha}_c$ is to use industry-level capital intensity as a proxy for occupation-level capital intensity: if the majority of employment in an occupation is employed in capital-intensive industries, it is considered a capital-intensive occupation.

Then, let Ω_j denote the set of occupations that are classified in task j, given in Table 1, and τ_c denote the employment share of occupation c in total employment. For each $j \in \{1, 2, 3\}$, we construct the measure of task-specific capital intensity α_j as a weighted sum of occupation-level capital intensities:

$$\alpha_j = \sum_{c \in \Omega_j} \frac{\tau_c}{\sum_{c' \in \Omega_j} \tau_{c'}} \times \hat{\alpha}_c.$$

Notice that the weight $\frac{\tau_c}{\sum_{c' \in \Omega_j} \tau_{c'}}$ is the employment share of occupation c in the total employment of the skill group to which it belongs.

The U.S. census is sufficient to obtain τ_c and $\tau_{(c,d)}$, and the data in 1980—the base year—are used. The data on τ_c and $\tau_{(c,d)}$, along with detailed descriptions, are reported in Appendix D.1. The measure of each κ_d is obtained from one minus the corresponding industry-level labor share measure, which is taken from the literature. The baseline labor share measures are taken from Lawrence (2015), in which the labor share is approximated by the labor compensation share. Two robustness checks are conducted using the two measures in Elsby et al. (2013), in which the payroll share is used to approximate the labor share. In particular, Lawrence (2015) provides labor share measures for 1987 and 2011; we use the former to construct $\hat{\alpha}_c$, as it is closer to the base year considered. Elsby et al. (2013) provide labor share measures for 1948, 1987, and 2011, and there are two measures for 1987: one is based on SIC codes, and the other is based on NAICS codes. We use both measures in 1987 for the robustness checks.

The baseline measures of capital intensities are reported in the first row of Table 2. It is evident that the middle-skill task is more capital intensive than the other two: the measured capital intensities are 0.391 for middle-skill, 0.333 for high-skill, and 0.336 for low-skill occupations. The results of the robustness check are reported in the second and third rows of Table 2. It is clear that the main conclusion of this subsection—that middle-skill occupations are more capital intensive than both high-and low-skill occupations—is robust across different measures of the labor share used in the literature.

Table 2: Capital Intensity Measures (1980)

Source for κ_d	High skill	Middle skill	Low skill
Lawrence (2015)	0.3328	0.3905	0.3355
Elsby et al. (2013), SIC	0.3104	0.4014	0.3222
Elsby et al. (2013), NAICS	0.2947	0.3660	0.3126

Table 3: Capital Intensity Measures (2010)

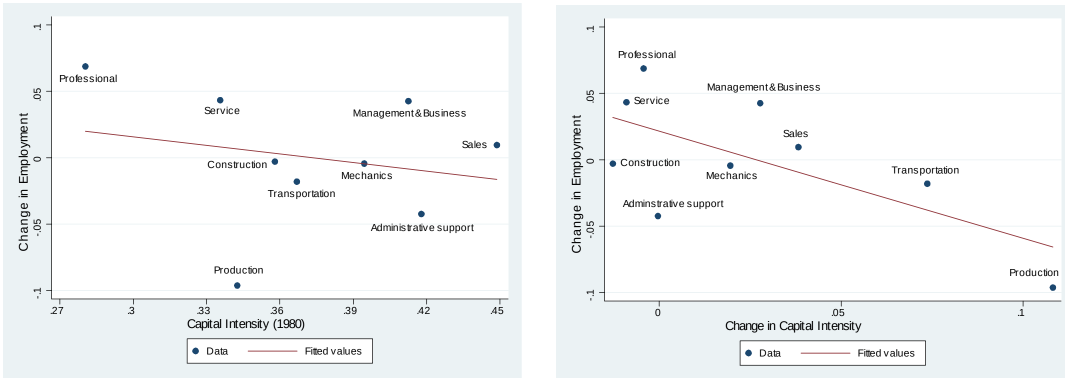
Source for κ_d	High skill	Middle skill	Low skill
Lawrence (2015)	0.3405	0.4341	0.3266
Elsby et al. (2013), NAICS	0.3022	0.4115	0.2914

5.3 Capital Intensity and Polarization

The second empirical contribution of this paper is to document how the occupation-level capital intensities change between 1980 and 2010 and how this change is related to job polarization. We construct the measures of occupational capital intensities for 2010 according to the methodology described above. The data come from the 2010 census, and the industry-level capital intensity is obtained using the 2011 labor share measures from Lawrence (2015) and Elsby et al. (2013). The results for the task-specific capital intensities are reported in Table 3. Clearly, the capital intensity of the middle-skill task increases significantly from 1980 to 2010, while the capital intensities of high-skill and low-skill tasks are nearly unchanged. In other words, the middle-skill task is not only more capital intensive, but it also becomes even more capital intensive, while the other two tasks remain at their previous capital intensity levels. Furthermore, as will be demonstrated below, empirical evidence suggests that job polarization is related not only to the higher capital intensity but also the increase in the capital intensity of middle-skill occupations.

In Figure 2, we plot the change in the employment share of each occupation against the respective 1980 capital intensity measures in the left panel. It presents a

Figure 2: Capital Intensity and Polarization



downward slope: the occupations with higher capital intensities experience more labor flowing out (or less labor flowing in) from 1980 to 2010. We also plot the change in the employment share of each occupation against the change in the respective capital intensity (the 2010 measure minus the 1980 measure) in the right panel. It also presents a prominent downward slope, suggesting that the increase in capital intensity is associated with the decline in the employment share of an occupation. Note that the change in the capital intensity can be represented as a form of labor-augmenting technological progress under the microfounded framework proposed by Jones (2005), and the labor-augmenting technological change is represented as better ideas that are found to make labor more productive.[11]

5.4 Discussion

Note that consulting the RBTC literature reveals a close connection between routineness and capital intensity: the tasks that are higher in measures of routineness are also the tasks that are more capital intensive. Based on this observation, this paper provides a different perspective on routinization, which is traditionally thought to be the engine that drives polarization. The early 1980s marks the starting point of job polarization, and in parallel to this labor market trend, the relative price of investment goods also experienced a significant decline. As documented in the

[11] Jones (2005) provides microfoundations for the Cobb-Douglas production function as a global production function derived from a series of ideas drawn from the Pareto distribution. The paper proposes that, to be consistent with balanced growth, the direction of technological change needs to be labor-augmenting.

literature, the relative price of investment goods has declined more rapidly since the early 1980s, suggesting an advancement in the technology of investment goods production. This paper thus proposes a new perspective arguing that part of the job polarization should be fundamentally attributed to the recent technological advances in producing investment goods. In other words, job polarization is not only due to the routine characteristic of middle-skill occupations and the hypothesis that there are technological advances biased toward routine tasks but is also caused by the capital-intensive characteristic of middle-skill occupations and the recent technological advances in producing investment goods.

Furthermore, as observed, an increase in the capital intensity of an occupation corresponds to a decrease in its employment share. A change in occupational capital intensity means a change in the production technology for the job task held by that occupation. An increase in capital intensity means that a task's production technology is shifted such that capital is used more heavily. Middle-skill task production experienced a noticeable increase in capital intensity, suggesting the existence of a technological change that is biased toward capital-intensive tasks and makes these tasks even more capital intensive. Not only is it possible for either a technological change in capital intensity or a decreasing price of investment goods alone to drive job polarization, but the coexistence of the two forces strengthens the effects of each force because of the interaction between them.

In the next section, we conduct a quantitative analysis to assess the extent to which these two forces—the technological advancement in investment goods production and the biased change in capital intensities—can explain job polarization; we also decompose the contribution of each force through counterfactual analysis.

6 Quantitative Analysis

The goal of the quantitative analysis is to assess how much of the employment polarization and wage polarization from 1980 to 2010 can be explained by the proposed hypothesis. In other words, we quantify the extent to which the changes in employment shares and relative wages can be explained by the change in the relative price of capital, given the observed heterogeneity in occupation-level capital intensities. Specifically, we choose 1980 as the base year and assume that all parameters except for ξ_x, r, and α_j are time-invariant. Regarding the calibration, for applicable parameters, we set the parameter values to match their empirical counterparts. The remaining parameters will be calibrated such that the model-implied moments fit the targeted data moments in the base year.

6.1 Calibration

First, we pin down the parameter values for which there are available estimates. The base-year task-level capital intensities are set at $(\alpha_1, \alpha_2, \alpha_3) = (0.3355, 0.3905, 0.3328)$ according to the estimates constructed in Subsection 5.2. Note that the measures proposed by Lawrence (2015) are used as the baseline since the labor compensation share is the preferred proxy for the labor share. The relative price of the investment good equals the inverse of the technology in the investment good sector $\frac{1}{\xi_x}$ in equilibrium, according to equation (10). Hence, ξ_x will be calibrated using estimates of the relative price of the investment good, which are taken from Justiniano et al. (2009). The authors show that the price exhibits a trend break in 1982; the two values of ξ_x for the calibration are obtained according to the two reported trends. We interpolate the value of ξ_x in 1980 using the pre-1982 trend; then, we extrapolate the value for 2010 using the post-1982 trend by assuming that the price trend continues until 2010.[12] The authors provide two estimates, one that uses the NIPA deflator for durable consumption and private investment and another that uses the GCV deflator for private equipment and software. Both estimates will be used in the quantitative analysis, while the GCV deflator will be considered as the baseline. To calibrate the capital rental rate R, measures of the depreciation rate and real interest rate are needed, according to equation (11). The depreciation rate δ is set at 0.1, following DeJong and Dave (2011), and data on the real interest rate are obtained from the World Bank.

The remaining parameters are calibrated by simulating the model to fit the base-year targets. Several assumptions are made. The task intensity parameters λ_j in the sectoral production function are assumed to be $\lambda_j = \frac{1}{3}$ for all j. We also normalize γ_1 and θ_3 to 1. This means that γ_2 and γ_3 measure the return to human capital of middle-skill and high-skill tasks relative to the low-skill task; similarly, θ_1 and θ_2 measure the returns on raw labor for low-skill and middle-skill tasks relative to the high-skill task. The workers' skills are assumed to follow $h \sim$ log-normal (μ, σ) with

[12]Their GCV estimate is available until 2000, and their NIPA estimate is available until 2009.

Table 4: Preset Parameters

Source	Parameter	Value	Note
Data	r	$\{0.057, 0.020\}$	$\{1980,2010\}$ values
	δ	0.1	
Normalization	γ_1	1	
	θ_3	1	
Assumption	$(\lambda_1, \lambda_2, \lambda_3)$	$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$	
Extrapolated	$1/\xi_x$	$\{0.741, 0.291\}$	$\{1980,2010\}$ values, GCV
from literature		$\{0.861, 0.419\}$	$\{1980,2010\}$ values, NIPA
Section 5.2	$(\alpha_1, \alpha_2, \alpha_3)$	(0.3355,0.3905,0.3328)	1980 value
		(0.3266,0.4341,0.3405)	2010 value

the mean normalized to 1.[13] The remaining six parameters to be calibrated are the parameters of the task production function γ_2, γ_3, θ_1, θ_2, the parameters of the skill distribution σ, and the elasticity of substitution of tasks η.

We simulate two equilibria, one for the base year 1980 and the other for 2010. First, we simulate the first equilibrium and find the six parameters such that the model moments fit the data moments in 1980, setting the investment good price $\frac{1}{\xi_x}$, real interest rate r, and capital intensities α_j at their 1980 values. Given all the other calibrated parameter values, we then simulate the other equilibrium, in which we let ξ_x, r, and α_j take their 2010 values while keeping all of the other parameters unchanged. The results of this simulation thus yield the model's predictions of employment shares and relative wages in 2010.

The base-year targets for the calibration are the four key moments of interest, the employment shares and relative average wages, and two additional moments, the income shares.[14] The preset parameters are summarized in Table 4. The calibrated parameter values are summarized in Table 5.

[13]The normalization implies the following restriction: $\mu + \frac{\sigma^2}{2} = 0$.

[14]The definition of income share and its empirical counterpart are presented in Appendix C.1.

Table 5: Calibrated Parameters

		Baseline (GCV)
Note	Parameter	Values
Task production	γ_2	5.33
	γ_3	57.55
	θ_1	273.32
	θ_2	9.22
Good production	η	0.0163
Skill distribution	σ	1.966

The table reports the calibrated parameter values using the relative price of the investment good based on the GCV deflator.

6.2 Model Fit

In the baseline quantitative analysis, we use the GCV-deflator-based measure of the relative price of the investment good, while we also conduct another analysis using the NIPA-deflator-based measure. The data and quantitative results are reported in Tables 6 and A.2, where $ES_j, \overline{w}_j, I_j$ denote the employment share, average wage, and income share of task j. We see that the model performs well at predicting both the declining employment share of middle-task occupations and the decreasing relative wage of middle-skill workers. When the capital rental rate declines (resulting from the decreasing price of the investment good and the decreasing real interest rate), labor flows out of middle-skill tasks into low-skill and high-skill tasks, generating employment polarization ($\Delta ES_2 < 0$ and $\Delta ES_{1,} \Delta ES_3 > 0$). Furthermore, the relative wages of middle-skill occupations to both low-skill and high-skill occupations decrease when the capital rental rate declines; this generates wage polarization (both $\Delta\frac{\overline{w}_2}{\overline{w}_1}$ and $\Delta\frac{\overline{w}_2}{\overline{w}_3} < 0$). Although the labor share is non-targeted, the model is able to capture its decreasing trend, as shown in the last rows of Tables 6 and A.2. The data indicate an approximately 6% decrease in the labor share, and the model predicts an approximately 3% decrease. Quantitatively, in the baseline quantitative analysis, the model is able to account for 54% of the decrease in employment in middle-skill occupations, as well as 39% and 94% of the increase in employment in high-skill and low-skill occupations, respectively. Regarding relative wages, 60% of the upper tail of wage polarization (i.e., the decline in the middle- to high-skill relative wage) can

Table 6: Model Fit: Baseline

	Data			Model			Model Fit
Variable	1980	2010	Δ of var.	1980	2010	Δ of var.	% explained
ES_3	25.83	36.95	11.12	25.82	30.12	4.30	38.67%
ES_2	61.35	45.89	-15.46	61.34	52.99	-8.35	54.01%
ES_1	12.83	17.16	4.33	12.83	16.89	4.06	93.76%
$\frac{\overline{w}_2}{\overline{w}_3}$	0.6519	0.4745	-27.2%	0.6517	0.5424	-16.8%	59.56%
$\frac{\overline{w}_2}{\overline{w}_1}$	1.6362	1.3720	-16.1%	1.6364	1.6296	-0.42%	2.58%
I_3	36.41	57.14	20.73	36.42	46.71	10.29	49.64%
I_2	56.38	33.68	-22.70	56.38	44.57	-11.81	52.03%
I_1	7.21	9.18	1.97	7.21	8.72	1.51	76.65%
Labor share	0.568	0.533	-6.16%	0.633	0.615	-2.84%	

The table reports the results using the relative price of the investment good based on the GCV deflator.

be explained, while only 2.5% of the lower tail of wage polarization is explained. The decomposition results are reported in Tables 7 and A.3. In the baseline model, the decreasing price of the investment good channel accounts for approximately 40% of the total fraction that this model can explain, and the change in the capital intensity channel accounts for the other 60%.

6.3 Discussion

In this subsection, we will explain the mechanism in greater detail by further utilizing the baseline quantitative analysis. The technological advancement in the production of the investment good increases the aggregate amount of capital in the economy. This is due to two effects. First, the price of the investment good decreases, and individuals' capital investment is increased. Second, the capital rental rate also decreases, and task producers' capital demand thus increases. The outputs of all tasks increase because aggregate capital increases, as capital is allocated to task production. How the additional capital should be allocated across tasks depends on the technologies of both final good production and task production. As noted, middle-skill task production is more capital intensive. From another perspective,

Table 7: Decomposition: Baseline

Variable	α_j invariant % explained	R invariant % explained	Decomposition: R contributes	Decomposition: α_j contributes
ES_3	12.05%	19.96%	41%	59%
ES_2	14.17%	27.49%	40%	60%
ES_1	19.86%	46.88%	39%	61%
$\frac{\overline{w}_2}{\overline{w}_3}$	19.49%	33.46%	39%	61%
$\frac{\overline{w}_2}{\overline{w}_1}$	0.62%	1.24%	41%	59%
I_3	14.95%	25.47%	41%	59%
I_2	14.98%	26.83%	40%	60%
I_1	14.72%	41.12%	36%	64%

The table reports the quantitative results using the relative price of the investment good based on the GCV deflator.

higher capital intensity also means that the production technology yields a higher return on capital. If the additional capital were allocated such that the new allocation of capital across tasks is as before, the output of the middle-skill task would grow more rapidly than that of the other two tasks. This would not be optimal because the three tasks are complementary in producing final goods. When tasks are complements, in the optimal capital allocation, the relative capital allocated to the middle-skill task must decrease. Since a smaller fraction of aggregate capital is used for middle-skill task production, it is optimal for labor to flow out. Therefore, this induces employment polarization: labor shifts away from middle-skill occupations and flows into high- and low-skill occupations.

A more detailed quantitative analysis, as summarized in Table 8, offers a better understanding of how this channel works. First, the table indicates that in middle-skill task production, the declining share of labor is in line with the declining share of capital. As mentioned previously, the middle-skill task employs a smaller share of capital when there is an increase in aggregate capital, which is a result of the interaction between two properties: the complementarity between different tasks and the higher return on capital of the middle-skill production technology. The share of labor employed in the middle-skill task accordingly also declines, which yields employment polarization. Table 8 also reports task-specific labor productivity, which is defined in Section 4.3, and the price of task output. Because higher capital intensity yields higher productivity if there is more aggregate capital, the price of the middle-skill task decreases while the prices of the high- and low-skill tasks increase. The labor productivity of the middle-skill task approximately triples, while the labor productivity of either the high- or low-skill task approximately doubles. The more rapid increase in labor productivity for the middle-skill task originates from both its higher capital intensity and more rapid increase in its capital intensity. In contrast to the RBTC literature, which assumes a task-specific technological shock that directly raises the productivity of middle-skill workers, this paper endogenously generates the increase in the labor productivity of middle-skill task production. In other words, as mentioned in Subsection 5.4, this paper argues that part of the RBTC that has been thought to drive job polarization should be fundamentally attributed to the interaction between the heterogeneity in task-specific capital intensity and the decrease in the capital rental rate.

Regarding the relative wages, Table 9 makes clear that the cutoffs $\left(\hat{h}_1, \hat{h}_2\right)$ shift inward at the new equilibrium. The middle-skill group loses the most able and the least able workers, and hence, its employment share shrinks. Owing to the property of the wage profile, as shown in equation (12), the movement of the slope b_j is in line with the movement of the intercept a_j. The quantitative results reported in Table 9 show that both a_1 and a_3 grow more rapidly than a_2, and that both b_1 and b_3 grow more rapidly than b_2. The low-skill and high-skill wages grow more rapidly than the middle-skill wage, in terms of both the returns on skills and the returns on their raw labor. This thus explains wage polarization.

6.4 Sensitivity Analysis

As mentioned in Subsection 6.2, we conducted a sensitivity analysis using the measure of the investment good price based on the NIPA deflator. Several additional sensitivity analysis exercises will be performed in this subsection.

6.4.1 Measurement of α_j

First, we construct another measure of occupation-level capital intensity by approximating the occupational labor share. First, we prove in Appendix C.1 that the model-implied task-level labor share is $1 - \alpha_j$. Then, we need to define several vari-

Table 8: Discussion: Employment Polarization

Variable	High skill 1980	High skill 2010	Middle skill 1980	Middle skill 2010	Low skill 1980	Low skill 2010
Share of labor	25.82	30.12	61.34	52.99	12.83	16.89
Share of capital	31.36	38.57	62.36	54.67	6.28	6.76
Task price	1.037	1.313	1.774	1.463	0.201	0.233
Labor productivity	100.6	218.1	42.0	123.7	208.0	399.9

Table 9: Discussion: Wage Polarization

	a_3	b_3	a_2	b_2	a_1	b_1	$\hat{h}_1$	$\hat{h}_2$
1980	1.19	68.47	27.36	18.04	27.64	0.10	0.016	0.519
2010	3.23	185.63	61.73	40.69	62.63	0.23	0.022	0.404
Δ	171.1%		125.6%		126.6%			

Table 8 and 9 report the quantitative results using the relative price of the investment good based on the GCV deflator.

ables before constructing the measure. We have 9 occupations and 13 industries; let c_m denote the occupations with $m \in \{1, ...9\}$, and let d_m denote the industries with $n \in \{1, ...13\}$. Let the employment and the total wages of occupation c_m be denoted by $L(c_m)$ and $W(c_m)$, respectively; we can then derive $W(c_m) = w_{c_m} L(c_m)$, where w_{c_m} is the average wage of occupation c_m. Let $L(c_m, d_n)$ denote the number of workers in occupation c_m in industry d_n, and let $W(c_m, d_n)$ and $Y(c_m, d_n)$ denote the sum of wages earned by and the sum of outputs produced by these workers, respectively. The industry-level labor share measure, denoted by $ls\,(d_n)$, is taken from the literature. The occupation-level labor share, denoted by $ls\,(c_m)$, is the objective to measure: $ls\,(c_m)$ is defined as

$$ls\,(c_m) = \frac{W(c_m)}{\sum_{n=1}^{13} Y(c_m, d_n)}.$$

There are no data available for $Y(c_m, d_n)$, and thus, we use the industry labor share measure to approximate it according to the following:

$$\frac{W(c_m, d_n)}{Y(c_m, d_n)} \approx ls(d_n),$$

where the underlying restriction is that $\frac{W(c_m,d_n)}{Y(c_m,d_n)}$ is identical for all occupations within each industry. Again, $W(c_m, d_n)$ can be approximated by $W(c_m, d_n) = w_{c_m} L(c_m, d_n)$. Given the above, we can rewrite $ls\,(c_m)$ as follows:

$$ls\,(c_m) \approx \frac{L(c_m)}{\sum_{n=1}^{13} \frac{L(c_m,d_n)}{ls(d_n)}}.$$

Equivalently,

$$ls\,(c_m) \approx \left[\sum_{n=1}^{13} \frac{\tau(c_m, d_n)}{ls(d_n)} \right]^{-1},$$

where $\tau(c_m, d_n)$ is the share of occupation c_m workers who are employed in industry d_n. For a comparison with Section 5.2, it is useful to adopt the same notation. Here, this measure of occupation-level capital intensity $\hat{\alpha}_c$ can be represented as follows:

$$\hat{\alpha}_c = 1 - \left[\sum_d \frac{\tau_{(c,d)}}{1 - \kappa_d} \right]^{-1}.$$

Similarly, let Ω_j denote the set of occupations that are classified in task j; τ_c denotes the employment share of occupation c in total employment. We construct the measure of task-specific capital intensity α_j for each $j \in \{1, 2, 3\}$ as follows:

$$\alpha_j = 1 - \left[\sum_{c \in \Omega_j} \left(\frac{\tau_c / \left(\sum_{c' \in \Omega_j} \tau_{c'} \right)}{1 - \hat{\alpha}_c} \right) \right]^{-1}$$

The results of this measure are reported in Table 10, where the measures of the industry-level labor share $ls(d_n)$ are taken from Lawrence (2015), as in the baseline. It is clear that the main properties of the baseline measure still hold under this new measure: first, the middle-skill task is more capital intensive; second, middle-skill capital intensity increases more than high-skill capital intensity does, while low-skill capital intensity decreases.

Table 10: Capital Intensity Measures (sensitivity analysis)

	High skill	Middle skill	Low skill
1980	0.3984	0.4505	0.3651
2010	0.4255	0.4910	0.3555

Source for κ_d: Lawrence (2015).

6.4.2 Task production function

While the model performs well at explaining the signs of the trends, the levels are underestimated. One explanation for the underestimation is that the task production technology is Cobb-Douglas. That is, the elasticity of substitution between capital and labor is unity. We have only allowed the capital intensity to vary across different tasks, but the elasticity of substitution has been kept the same under the Cobb-Douglas assumption. It is possible that not only is middle-skill task production more capital intensive, but middle-skill workers are also more easily substituted by physical capital. Specifically, job tasks are produced by CES technology in which the capital-labor elasticity of substitution is heterogeneous:

$$y_j = \left\{ \left[(1-\alpha_j) \left(\gamma_j h + \theta_j \right) \right]^{\nu_j} + \left[\alpha_j k \right]^{\nu_j} \right\}^{\frac{1}{\nu_j}},$$

and it is assumed that $\nu_2 > \nu_1$ and $\nu_2 > \nu_3$. When middle-skill labor is more substitutable by capital, it is a natural prediction that the effect of a declining price of capital goods on job polarization will be strengthened.

6.5 Comparison with the Literature

Finally, we compare our model with the RBTC hypothesis in terms of its quantitative contributions. Although RBTC is modeled in various ways, as mentioned in the literature review, these models share a common characteristic, namely, that RBTC directly increases middle-skill workers' productivity. This direct effect holds true regardless of whether job polarization is present, as opposed to our paper, in which the increase in middle-skill workers' labor productivity occurs alongside job polarization. To capture this feature, we incorporate the RBTC channel in a compatible way: we assume that routinization takes the form of factor-augmenting technological change that makes middle-skill labor more productive or, in other words, increases the return on raw labor.

The idea of the comparison exercise is to conduct a counterfactual analysis such that the RBTC channel can account for as much of ΔES_2 as this model does. Let A_j denote the labor-augmenting technology of task j, and the task production function is as follows:

$$y_j = \left(\gamma_j h + A_j \theta_j\right)^{1-\alpha_j} k^{\alpha_j},$$

and RBTC is represented by A_2 increasing over time while A_1 and A_3 are constant. Routinization is associated with the process of computerization or automation, which replaces workers' routine skills but enhances productivity per worker. It could thus be modeled as the technological change that is labor-augmenting and biased toward the raw labor in middle-skill tasks.

We conduct a counterfactual analysis, in which let A_2 increase to the extent that the model-implied employment share of middle-skill workers decreases to 52.99, the extent that the channel proposed by this paper can explain.[15] The results are shown in Table 11. The RBTC channel performs better at the upper tail of employment polarization but worse at the lower tail. Regarding the relative wages, the RBTC channel overshoots both tails of wage polarization. The calibrated A_2 is 2.88, suggesting that routinization needs to boost the relative productivity of middle-skill workers' raw labor by approximately three times to explain the decrease in the middle-skill employment share to the same extent as this paper can. Regarding the labor share, however, the RBTC channel prediction is of the opposite sign of the values in the data. The counterfactual analysis shows that, in the presence of RBTC, the labor share would increase. This result contradicts the data, in which the labor share decreases.

Table 12 shows that the RBTC channel also results in a decreasing share of capital in middle-skill task production, which is associated with the outflow of employment away from middle-skill occupations. The movements of task-specific prices and labor productivity are also similar: for the middle-skill task, the output price decreases, and there is a 70% increase in the labor productivity; for high- and low-skill tasks, the output prices increase, and there is a 20% increase in the labor produc-

[15]Note that the channel proposed by this paper is shut down in the counterfactual analysis.

tivity. Table 13 shows, however, that the inward movement of the two skill thresholds originates from a different mechanism. While the slopes and the intercepts of both the high-skill and low-skill wage profiles increase, which is the same as in this model, the slope of the middle-skill wage profile decreases, and the intercept increases. This means that, as RBTC replaces middle-skill workers' skills, the wage paid for the skill component decreases in the presence of RBTC, while the wage paid for the raw labor component still increases.

6.6 Policy Experiments

In this subsection, we conduct policy experiments in the context of education reform to study the effects on job polarization. We consider two types of policies. The first is a program to improve education quality, whereby overall human capital is increased while its variation is unchanged. The second is the heterogeneous grouping program, whereby the variation of human capital is decreased while its mean is preserved. Since the focus is on how the policy impacts the employment distribution and relative wages, this analysis abstracts from how the policy is financed.

In the first policy experiment, the program to improve education quality, we let the mean of workers' skills $E(h)$ increase from 1 to 1.5 while keeping the variance $V(h)$ unchanged. A policy that improves education quality would increase the overall levels of human capital but not affect its variance, and hence, it can be represented by a rightward shift of the skill distribution. In the second policy experiment, the heterogeneous grouping program, we let $V(h)$ decrease to half of its previous level while keeping $E(h)$ unchanged. The heterogeneous grouping would make students' skills more concentrated, and thus, it can be represented by a mean-preserving concentration of the skill distribution. The results of the policy experiments are reported in Table 14.

The introduction of the program to improve education quality would increase the relative wages of middle-skill workers with respect to both high- and low-skill workers, the opposite of the pattern observed under wage polarization. There is nearly no change in b_3 or a_3, while b_2 and a_2 increase by 80%, and b_1 and a_1 nearly double. This is because when the skill distribution shifts rightward, the workers in the right tail of the distribution benefit the least, while the workers in the middle and bottom of the distribution enjoy a considerable improvement. The increase in $\frac{\overline{w}_2}{\overline{w}_3}$ is thus explained by this property. $\frac{\overline{w}_2}{\overline{w}_1}$ also increases because, compared with the

Table 11: Counterfactual: RBTC

Variable	Data 1980	Data 2010	Data Δ of var.	Model 1980	Model 2010	Model Δ of var.	Model Fit % explained
ES_3	25.83	36.95	11.12	25.82	31.09	5.27	47.39%
ES_2	61.35	45.89	-15.46	61.34	52.99	-8.35	54.01%
ES_1	12.83	17.16	4.33	12.83	15.92	3.09	71.36%
$\frac{\overline{w}_2}{\overline{w}_3}$	0.6519	0.4745	-27.2%	0.6517	0.4438	-31.9%	117.28%
$\frac{\overline{w}_2}{\overline{w}_1}$	1.6362	1.3720	-16.1%	1.6364	1.2186	-25.5%	158.39%
I_3	36.41	57.14	20.73	36.42	51.48	15.06	30.40%
I_2	56.38	33.68	-22.70	56.38	38.93	-17.45	76.87%
I_1	7.21	9.18	1.97	7.21	9.60	2.39	121.32%
Labor share	0.568	0.533	-6.16%	0.633	0.643	1.58%	

The table reports the results using the relative price of the investment good based on the GCV deflator.

Table 12: RBTC Counterfactual: Employment Polarization

Variable	High skill 1980	High skill 2010	Middle skill 1980	Middle skill 2010	Low skill 1980	Low skill 2010
Share of labor	25.82	31.09	61.34	52.99	12.83	15.92
Share of capital	31.36	46.27	62.36	44.99	6.28	8.74
Task price	1.037	1.498	1.774	1.237	0.201	0.273
Labor productivity	100.6	120.9	42.0	71.1	208.0	242.8

Table 13: RBTC Counterfactual: Wage Polarization

	a_3	b_3	a_2	b_2	a_1	b_1	$\hat{h}_1$	$\hat{h}_2$
1980	1.19	68.47	27.36	18.04	27.64	0.10	0.016	0.519
2010	2.06	118.76	43.65	9.99	43.85	0.16	0.020	0.382
Δ	173.4%		159.5%	55.4%	158.7%			

Table 12 and 13 report the counterfactual analysis results using the relative price of the investment good based on the GCV deflator.

Table 14: Policy Experiments

	Baseline			$E(h)' = 1.5$		$V(h)' = \frac{1}{2}V(h)$	
Variable	1980	2010	Δ of var.	2010	Δ of var.	2010	Δ of var.
ES_3	25.82	30.12	4.30	28.77	2.95	31.24	5.42
ES_2	61.34	52.99	-8.35	51.39	-9.95	52.09	-9.28
ES_1	12.83	16.89	4.06	19.85	7.02	16.67	3.84
$\frac{\overline{w}_2}{\overline{w}_3}$	0.6517	0.5424	-16.8%	0.7341	12.6%	0.6218	-4.6%
$\frac{\overline{w}_2}{\overline{w}_1}$	1.6364	1.6296	-0.42%	1.8877	15.4%	1.6149	-1.3%
b_3	68.47	185.63	171%	137.28	101%	172.19	152%
b_2	18.04	40.69	126%	50.68	181%	43.28	140%
b_1	0.10	0.23	127%	0.29	195%	0.25	146%
a_3	1.19	3.23	171%	2.39	101%	2.99	151%
a_2	27.36	61.73	126%	76.89	181%	65.66	140%
a_1	27.64	62.63	127%	80.56	192%	67.21	143%
$\hat{h}_1$	0.016	0.022		0.860		0.486	
$\hat{h}_2$	0.519	0.404		0.073		0.036	

The table reports the results using the relative price of the investment good based on the GCV deflator.

baseline, there are more workers flowing into the low-skill task, which drives down the wage of low-skill workers.

The introduction of the heterogeneous grouping program would dampen the upper tail of wage polarization but strengthen the lower tail of wage polarization. The result that $\frac{\overline{w}_2}{\overline{w}_3}$ decreases less is because the increases in b_3 and a_3 are not as large as in the baseline, while the increases in b_2 and a_2 as well as b_1 and a_1 are greater than in the baseline. The concentration of the skill distribution results in fewer top earners, which decreases the average wage of high-skill workers and thus makes $\frac{\overline{w}_2}{\overline{w}_3}$ decrease less. The larger decrease in $\frac{\overline{w}_2}{\overline{w}_1}$ is because, compared with the baseline, fewer workers are flowing into the low-skill task, which raises the wage of low-skill workers.

7 Conclusion

This paper proposes a new perspective to explain the prevalent job polarization experienced over the past few decades and demonstrates that the proposed mechanism can be directly supported by empirical evidence. This paper contributes to the literature both empirically and theoretically. We document the heterogeneity in job task production among different skill groups by constructing a measure of occupation-level capital intensity. This paper shows that middle-skill tasks are more capital intensive and that the capital intensity of such tasks has been growing more substantially. This feature—in addition to the conventional wisdom that middle-skill tasks are more routine—provides a new mechanism to explain job polarization. We construct a task-based model in which there is endogenous positive sorting between workers' skills and tasks' production technologies. The model predicts that the recent decline in the relative price of capital goods can induce both employment and wage polarization. The increase in the capital intensity of the middle-skill task can also induce job polarization. Moreover, alongside job polarization, the labor productivity of middle-skill workers increase relative to high- and low-skill workers.

While the literature focuses on the role of computer capital, our paper argues that equipment capital should also be responsible for explaining job polarization. Investment technological progress has been considered an important factor in driving several labor market phenomena, such as the declining labor share and the rising skill premium. Our paper aims to address its importance in accounting for job polarization, and we thus argue that these macroeconomic labor trends that have been found to be salient since 1980s are closely related.

A useful extension of this paper would be to develop a nested model with both our capital intensity channel and the RBTC channel built in. In our quantitative exercise on the comparison with RBTC in Subsection 6.5, our setting can be viewed as a reduced-form approximation of the existing settings in the literature. Although the existing settings of RBTC can all explain job polarization, there are noteworthy differences in their fundamental meanings.[16] It will be valuable to distinguish their key differences in terms of their ability to explain other important patterns of interest. Having a justifiable RBTC setting and then constructing a rich model that integrates both our channel and the RBTC channel will enable us to evaluate the sources of job polarization in a meaningful way.

Finally, as mentioned earlier, the engine in our paper to explain job polarization—the positive shock to the investment technology—has also been believed to give rise to other prominent labor market trends (e.g., Krusell et al. [2000] and Karabarbounis and Neiman [2014]). It would be an interesting extension to construct a unified framework that can jointly explain these long-run labor market patterns. Moreover, we have documented that the capital intensity of middle-skill task production has increased more rapidly. While we consider it to be an exogenous process in the model, we offer a possible explanation for it. This increase might result from a directed technological change induced by the investment technology shock. That is, the capital-intensive feature of the middle-skill task makes its producers more inclined to adjust the production technology, given the decreasing price of capital. Following Jones (2005), we view the increase in capital intensity as the arrival of labor-augmenting technological progress, and the adoption of a new type of capital or a rearrangement of capital-related production procedures could enhance the efficiency of middle-skill workers (referred to as labor-enhancing technologies). Though not the main focus in our paper, it might be of interest in future research to endogenize this directed technological change under a framework similar to those of León-Ledesma and Satchi (2017) and Uras and Wang (2017) to provide microfounded explanations for the change in task-specific capital intensity.

[16]See the literature review for further details.

Appendix

Appendix A Derivations

A.1 Relative Prices of Tasks

Given the price of job task q_j, cost-minimization implies that consumption good producers' demand for task j is $Y_{c,j} = \lambda_j \left[\frac{q_j}{p_c}\right]^{-\eta} C$, and investment good producers' demand is $Y_{x,j} = \lambda_j \left[\frac{q_j}{p_x}\right]^{-\eta} X$. It can thus be shown that for all j, $\frac{Y_{c,j}}{Y_{x,j}} = \left(\frac{p_c}{p_x}\right)^{\eta} \frac{C}{X}$ holds. Define s_j as the fraction of Y_j being allocated to the consumption good sector and $(1-s_j)$ the fraction allocated to the investment good sector; i.e., $Y_{c,j} = s_j Y_j$ and $Y_{x,j} = (1-s_j)Y_j$. From the above, we know that $\frac{Y_{c,1}}{Y_{x,1}} = \frac{Y_{c,2}}{Y_{x,2}} = \frac{Y_{c,3}}{Y_{x,3}}$ and thus obtain

$$\frac{s_1}{1-s_1} = \frac{s_2}{1-s_2} = \frac{s_3}{1-s_3},$$

which jointly imply that $s_j = s$ for all $j \in J$. That is,

$$Y_{c,j} = sY_j \quad \text{and} \quad Y_{x,j} = (1-s)\,Y_j \quad \text{for all } j \in J. \tag{16}$$

This property states that tasks are allocated in the same way to final good production: for each task, the same fraction of each task's output, denoted s, is used to produce the consumption good, and the other fraction $(1-s)$ is used to produce the investment good. This property immediately implies the following conditions:

$$\frac{Y_{c,j}}{Y_{c,j'}} = \frac{Y_{x,j}}{Y_{c,j\prime}} = \frac{Y_j}{Y_{j\prime}} \quad \forall j, j' \in J.$$

By applying the above conditions to the task demand functions, we have

$$\begin{aligned} Y_{c,j} &= \lambda_j q_j^{-\eta} C, \\ Y_{x,j} &= \lambda_j q_j{}^{-\eta} \xi_x^{-1} X. \end{aligned}$$

We thus obtain the equilibrium conditions for the relative prices of tasks:

$$\frac{q_j}{q_{j'}} = \left(\frac{\lambda_j}{\lambda_{j'}}\right)^{\frac{1}{\eta}} \left(\frac{Y_j}{Y_{j'}}\right)^{-\frac{1}{\eta}} \quad \forall j, j' \in J.$$

A.2 Wage Profile

By solving the task producers' maximization problem (7), we obtain their policy function for capital $k_j(h)$ and the equilibrium wage profile $w_j(h)$. The first-order conditions for task producers' maximization problem with respect to k and h are as follows, respectively:

$$q_j \alpha_j \left(\gamma_j h + \theta_j\right)^{1-\alpha_j} k_j(h)^{\alpha_j - 1} - \xi_x^{-1}(r+\delta) = 0, \tag{17}$$

$$q_j(1-\alpha_j)\gamma_j \left(\frac{k_j(h)}{\gamma_j h + \theta_j}\right)^{\alpha_j} + q_j \alpha_j \left(\frac{k_j(h)}{\gamma_j h + \theta_j}\right)^{\alpha_j - 1} k_j'(h) \\ - w_j'(h) - \xi_x^{-1}(r+\delta) k_j'(h) = 0. \tag{18}$$

By rearranging (17), we obtain the policy function for capital $k_j(h)$ of a task-j producer who hires a skill-h worker:

$$k_j(h) = \left[q_j \frac{\alpha_j \xi_x}{(r+\delta)}\right]^{\frac{1}{1-\alpha_j}} \left(\gamma_j h + \theta_j\right), \tag{19}$$

and its first derivative is

$$k_j'(h) = \gamma_j \left[q_j \frac{\alpha_j \xi_x}{(r+\delta)}\right]^{\frac{1}{1-\alpha_j}}. \tag{20}$$

By combining (19) and (20) with (18), we obtain that the first derivative of the wage profile $w_j(h)$ satisfies

$$w_j'(h) = \gamma_j q_j^{\frac{1}{1-\alpha_j}} \xi_x^{\frac{\alpha_j}{1-\alpha_j}} (1-\alpha_j) \left(\frac{\alpha_j}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}},$$

and the second derivative satisfies $w_j''(h) = 0$. Thus, it can be inferred that the wage profile $w_j(h)$ is a linear function of h. That is, $w_j(h)$ can be expressed as

$$w_j(h) = a_j + b_j h.$$

Its slope, equal to $w_j'(h)$, is thus

$$b_j = \gamma_j q_j^{\frac{1}{1-\alpha_j}} \xi_x^{\frac{\alpha_j}{1-\alpha_j}} (1-\alpha_j) \left(\frac{\alpha_j}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}}.$$

Since there is free entry and exit of task producers in the competitive market, the zero-profit condition always holds in equilibrium. Specifically, after entry, a producer in task j hiring a worker with skill level h earns

$$\gamma_j q_j^{\frac{1}{1-\alpha_j}} \xi_x^{\frac{\alpha_j}{1-\alpha_j}} (1-\alpha_j) \left(\frac{\alpha_j}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}} \left(\gamma_j h + \theta_j\right) - (a_j + b_j h).$$

By applying the zero-profit condition, we can then solve for the intercept of $w_j(h)$:

$$a_j = \theta_j \gamma_j q_j^{\frac{1}{1-\alpha_j}} \xi_x^{\frac{\alpha_j}{1-\alpha_j}} (1-\alpha_j) \left(\frac{\alpha_j}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}}.$$

A.3 Market Clearing Conditions

Goods market

We rewrite equation (5) by plugging in (19):

$$y_j(h) = \left(\frac{q_j \alpha_j \xi_x}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}} \left(\gamma_j h + \theta_j\right), \tag{21}$$

and the aggregate output of task j is thus

$$Y_j = \int_{\mathbb{O}_j} y_j(h)\, dF(h) = \left(\frac{q_j \alpha_j \xi_x}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}} \left(\gamma_j \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_j\right) ES_j.$$

The left-hand side of the goods market clearing condition (9) equals

$$C + p_x X = \left[\sum_{j\in J} \lambda_j^{\frac{1}{\eta}} Y_{c,j}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} + \left[\sum_{j\in J} \lambda_j^{\frac{1}{\eta}} Y_{x,j}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} = \left[\sum_{j\in J} \lambda_j^{\frac{1}{\eta}} Y_j^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}},$$

where the first and the second equalities are derived by applying (11) and (16), respectively. For the right-hand side of (9), we derive the first argument by plugging in (12):

$$\sum_{j\in J} \int_{\mathbb{O}_j} w_j(h)\, dF(h) = \sum_{j\in J} q_j^{\frac{1}{1-\alpha_j}} \left(1-\alpha_j\right) \left(\frac{\alpha_j \xi_x}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}} \left(\gamma_j \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_j\right) ES_j.$$

For the second argument, the aggregate demand for capital is the summation over task producers' capital demand (19):

$$K^D = \sum_{j\in J} \int_{\mathbb{O}_j} k_j(h)\, dF(h) = \sum_{j\in J} \left(\frac{q_j \alpha_j \xi_x}{r+\delta}\right)^{\frac{1}{1-\alpha_j}} \left(\gamma_j \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_j\right) ES_j.$$

We thus rewrite the goods market clearing condition (9) as follows:

$$
\begin{aligned}
& \left[\sum_{j\in J}\lambda_j^{\frac{1}{\eta}}\left(\left(\frac{q_j\alpha_j\xi_x}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}}\left(\gamma_j\exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_j\right)ES_j\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} \\
= & \sum_{j\in J}q_j^{\frac{1}{1-\alpha_j}}\left(\frac{\alpha_j\xi_x}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}}\left(\gamma_j\exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_j\right)ES_j.
\end{aligned}
$$

Capital market

In steady state, the aggregate capital supply is $K^S = \frac{1}{\delta}X$. From (16), we know that $Y_{x,j} = (1-s)\,Y_j$ for all j. Therefore, we derive

$$
\begin{aligned}
K^S &= \frac{1}{\delta}\xi_x\left[\sum_{j\in J}\lambda_j^{\frac{1}{\eta}}Y_{x,j}^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} = \frac{1}{\delta}\xi_x(1-s)\left[\sum_{j\in J}\lambda_j^{\frac{1}{\eta}}Y_j^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} \\
&= \frac{1}{\delta}\xi_x(1-s)\left[\sum_{j\in J}\lambda_j^{\frac{1}{\eta}}\left(\left[q_j\frac{\alpha_j\xi_x}{(r+\delta)}\right]^{\frac{\alpha_j}{1-\alpha_j}}\left(\gamma_j\exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_j\right)ES_j\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}.
\end{aligned}
$$

Equating K^D and K^S yields the capital market clearing condition:

$$
\begin{aligned}
& \sum_{j\in J}\xi_x^{\frac{\alpha_j}{1-\alpha_j}}\left(\frac{q_j\alpha_j}{r+\delta}\right)^{\frac{1}{1-\alpha_j}}\left(\gamma_j\exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_j\right)ES_j \\
= & \frac{1}{\delta}(1-s)\left[\sum_{j\in J}\lambda_j^{\frac{1}{\eta}}\left(\left[q_j\frac{\alpha_j\xi_x}{(r+\delta)}\right]^{\frac{\alpha_j}{1-\alpha_j}}\left(\gamma_j\exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_j\right)ES_j\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}.
\end{aligned}
$$

Appendix B Proofs

B.1 Proof of Proposition 1 (Positive Sorting)

First, we prove the following.

Lemma 1 *There exist $\underline{v}, \overline{v}$ such that $\frac{\varpi_2(q_2,\xi_x)}{\varpi_j(q_j,\xi_x)}$ is bounded by $[\underline{v}, \overline{v}]$ for all $\xi_x \in \left[\underline{\xi}, \overline{\xi}\right]$ and $j \in \{1,3\}$.*

Proof. See B.1.1. ■

Since the wage profile is linear in h, positive sorting can be ensured if both $a_1 > a_2 > a_3$ and $b_3 > b_2 > b_1$ hold, which in turn requires the following conditions to hold:

$$\frac{\gamma_3}{\gamma_2} > \frac{\varpi_2(q_2,\xi_x)}{\varpi_3(q_3,\xi_x)} > \frac{\theta_3}{\theta_2},$$
$$\frac{\theta_1}{\theta_2} > \frac{\varpi_2(q_2,\xi_x)}{\varpi_1(q_1,\xi_x)} > \frac{\gamma_1}{\gamma_2}.$$

Given Assumption 3, we can rewrite the above conditions as:

$$1+\varepsilon_2^{\gamma} > \bar{v} \quad \text{and} \quad 1+\varepsilon_1^{\theta} > \bar{v},$$
$$\frac{1}{1+\varepsilon_1^{\gamma}} < \underline{v} \quad \text{and} \quad \frac{1}{1+\varepsilon_2^{\theta}} < \underline{v}.$$

Thereafter, as long as $\varepsilon_2^{\gamma}, \varepsilon_1^{\theta} > \bar{v}-1$ and $\varepsilon_1^{\gamma}, \varepsilon_2^{\theta} > \frac{1}{\underline{v}} - 1$, then the slope of the wage profile b_j, is increasing in j, and the intercept a_j is decreasing in j. To conclude Proposition 1, we set

$$\underline{\varepsilon} = \max\left\{\overline{v}-1, \frac{1}{\underline{v}}-1\right\}.$$

QED

B.1.1 Proof of Lemma 1

Given the definition of $\varpi_j(q_j,\xi_x)$ in Subsection 3.3, we can derive the following:

$$\frac{\varpi_2(q_2,\xi_x)}{\varpi_j(q_j,\xi_x)} = \frac{q_2^{\frac{1}{1-\alpha_2}}}{q_j^{\frac{1}{1-\alpha_j}}} \xi_x^{\frac{\alpha_2}{1-\alpha_2}-\frac{\alpha_j}{1-\alpha_j}} \frac{(\alpha_2-1)\,\alpha_2^{\frac{\alpha_2}{1-\alpha_2}}}{(\alpha_j-1)\,\alpha_j^{\frac{\alpha_j}{1-\alpha_j}}} (r+\delta)^{-\frac{\alpha_2}{1-\alpha_2}+\frac{\alpha_j}{1-\alpha_j}}. \tag{22}$$

Moreover, as will be shown in the proof of Proposition 2, $q_2^{\frac{1}{1-\alpha_2}}/q_j^{\frac{1}{1-\alpha_j}}$ is negatively correlated with ξ_x, given $\alpha_2 > \alpha_j$ for $j = 1,3$. To ease notation, we let $Q_{2j}(\xi_x)$ denote $q_2^{\frac{1}{1-\alpha_2}}/q_j^{\frac{1}{1-\alpha_j}}$ henceforth. Since $Q_{2j}(\xi_x)$ is continuous and decreasing in ξ_x, we can thus express it as $Q_{2j}(\xi_x) = Q_{2j}\left(\underline{\xi}\right) - \tilde{\rho}_{j,\xi_x}\left(\xi_x - \underline{\xi}\right)$, where $\tilde{\rho}_{j,\xi_x}$ is a positive scalar with its value depending on ξ_x. Accordingly, we can derive the following:

$$\begin{aligned}&\frac{\partial}{\partial \xi_x} Q_{2j}(\xi_x)\, \xi_x^{\frac{\alpha_2}{1-\alpha_2}-\frac{\alpha_j}{1-\alpha_j}} \\ = \;& \xi_x^{\frac{2\alpha_2-\alpha_2\alpha_j-1}{(1-\alpha_2)(1-\alpha_j)}} \left\{ \left(\frac{\alpha_j}{1-\alpha_j} - \frac{\alpha_2}{1-\alpha_2} - 1 \right) \xi_x + \left(\frac{\alpha_2}{1-\alpha_2} - \frac{\alpha_j}{1-\alpha_j} \right) \left(\tilde{\rho}_{j,\xi_x}\underline{\xi} + Q_{2j}\left(\underline{\xi}\right) \right) \right\},\end{aligned}$$

and it is clear that it is bounded for all $\xi_x \in \left[\underline{\xi}, \overline{\xi}\right]$. Furthermore, since $2\alpha_2 - \alpha_2\alpha_j - 1 < 0$ if $\alpha_2 < \frac{1}{2}$ and $\frac{\alpha_j}{1-\alpha_j} - \frac{\alpha_2}{1-\alpha_2} - 1 < 0$ as $\alpha_j < \alpha_2$, the above is bounded for all ξ_x if $\alpha_2 < \frac{1}{2}$. Hence, it immediately implies that (22) is bounded—and we denote that (22) is bounded by $[\underline{v}, \overline{v}]$. *QED*

B.2 Proof of Proposition 2 (Employment Polarization)

We use two sets of conditions to complete the proof: first, the conditions of the relative prices of tasks and, second, the conditions of the wage profile at the skill thresholds.

First, given (8), (13), and that $h \sim \text{log-normal}(\mu, \sigma)$, we can rewrite the conditions of relative prices of tasks (14) as

$$\frac{ES_j}{ES_{j'}} \frac{q_j^{\eta+\frac{\alpha_j}{1-\alpha_j}}}{q_{j'}^{\eta+\frac{\alpha_{j'}}{1-\alpha_{j'}}}} = \xi_x^{-\frac{\alpha_j}{1-\alpha_j}+\frac{\alpha_{j'}}{1-\alpha_{j'}}} (r+\delta)^{\frac{\alpha_j}{1-\alpha_j}-\frac{\alpha_{j'}}{1-\alpha_{j'}}} \frac{\lambda_j \alpha_j^{\frac{-\alpha_j}{1-\alpha_j}} \left(\gamma_j \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_j \right)^{-1}}{\lambda_{j'} \alpha_{j'}^{\frac{-\alpha_{j'}}{1-\alpha_{j'}}} \left(\gamma_{j'} \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_{j'} \right)^{-1}} \tag{23}$$

for all $j, j' \in J$. To facilitate the analysis, we assign $j = 2$ and $j' = 1,3$. Given Assumption 2 and thus $\frac{\alpha_2}{1-\alpha_2} > \frac{\alpha_{j'}}{1-\alpha_{j'}}$ for $j' = 1,3$, we can see that an increase in ξ_x

results in a decrease in the left-hand side of (23):

$$\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\frac{q_2^{\eta+\frac{\alpha_2}{1-\alpha_2}}}{q_1^{\eta+\frac{\alpha_1}{1-\alpha_1}}}\right) < 0, \tag{24}$$

$$\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_3}\frac{q_2^{\eta+\frac{\alpha_2}{1-\alpha_3}}}{q_3^{\eta+\frac{\alpha_1}{1-\alpha_3}}}\right) < 0. \tag{25}$$

Or equivalently,

$$\frac{\partial}{\partial \xi_x}\left(\log\frac{ES_2}{ES_1}+\log\frac{q_2^{\eta+\frac{\alpha_2}{1-\alpha_2}}}{q_1^{\eta+\frac{\alpha_1}{1-\alpha_1}}}\right) < 0, \tag{P2-1}$$

$$\frac{\partial}{\partial \xi_x}\left(\log\frac{ES_2}{ES_3}+\log\frac{q_2^{\eta+\frac{\alpha_2}{1-\alpha_3}}}{q_3^{\eta+\frac{\alpha_1}{1-\alpha_3}}}\right) < 0. \tag{P2-2}$$

Second, at skill thresholds $(\hat{h}_1, \hat{h}_2)$, the following conditions must hold:

$$w_1\left(\hat{h}_1\right)=w_2\left(\hat{h}_1\right);\ \ w_2\left(\hat{h}_2\right)=w_3\left(\hat{h}_2\right).$$

Given the wage profile (12) and after some rearrangement, we derive the following conditions:

$$\frac{q_2^{\frac{1}{1-\alpha_2}}\left(\gamma_2\hat{h}_1+\theta_2\right)}{q_1^{\frac{1}{1-\alpha_1}}\left(\gamma_1\hat{h}_1+\theta_1\right)} = \xi_x^{-\frac{\alpha_2}{1-\alpha_2}+\frac{\alpha_1}{1-\alpha_1}}\left(r+\delta\right)^{\frac{\alpha_2}{1-\alpha_2}-\frac{\alpha_1}{1-\alpha_1}}\frac{\left(1-\alpha_1\right)\alpha_1^{\frac{\alpha_1}{1-\alpha_1}}}{\left(1-\alpha_2\right)\alpha_2^{\frac{\alpha_2}{1-\alpha_2}}}, \tag{26}$$

$$\frac{q_2^{\frac{1}{1-\alpha_2}}\left(\gamma_2\hat{h}_2+\theta_2\right)}{q_3^{\frac{1}{1-\alpha_3}}\left(\gamma_3\hat{h}_2+\theta_3\right)} = \xi_x^{-\frac{\alpha_2}{1-\alpha_2}+\frac{\alpha_3}{1-\alpha_3}}\left(r+\delta\right)^{\frac{\alpha_2}{1-\alpha_2}-\frac{\alpha_3}{1-\alpha_3}}\frac{\left(1-\alpha_3\right)\alpha_3^{\frac{\alpha_3}{1-\alpha_3}}}{\left(1-\alpha_2\right)\alpha_2^{\frac{\alpha_2}{1-\alpha_2}}}. \tag{27}$$

Similarly, an increase in ξ_x results in a decrease in the left-hand side of (26) and (27):

$$\frac{\partial}{\partial \xi_x}\left[\frac{q_2^{\frac{1}{1-\alpha_2}}\left(\gamma_2 \hat{h}_1+\theta_2\right)}{q_1^{\frac{1}{1-\alpha_1}}\left(\gamma_1 \hat{h}_1+\theta_1\right)}\right] < 0, \tag{28}$$

$$\frac{\partial}{\partial \xi_x}\left[\frac{q_2^{\frac{1}{1-\alpha_2}}\left(\gamma_2 \hat{h}_2+\theta_2\right)}{q_3^{\frac{1}{1-\alpha_3}}\left(\gamma_3 \hat{h}_2+\theta_3\right)}\right] < 0. \tag{29}$$

Or equivalently,

$$\frac{\partial}{\partial \xi_x}\left(\log \frac{q_2^{\frac{1}{1-\alpha_2}}}{q_1^{\frac{1}{1-\alpha_1}}}+\log \frac{\gamma_2 \hat{h}_1+\theta_2}{\gamma_1 \hat{h}_1+\theta_1}\right) < 0, \tag{P2-3}$$

$$\frac{\partial}{\partial \xi_x}\left(\log \frac{q_2^{\frac{1}{1-\alpha_2}}}{q_3^{\frac{1}{1-\alpha_3}}}+\log \frac{\gamma_2 \hat{h}_2+\theta_2}{\gamma_3 \hat{h}_2+\theta_3}\right) < 0. \tag{P2-4}$$

Finally, by substituting $\xi_x^{-\frac{\alpha_2}{1-\alpha_2}+\frac{\alpha_1}{1-\alpha_1}}$ and $\xi_x^{-\frac{\alpha_2}{1-\alpha_2}+\frac{\alpha_3}{1-\alpha_3}}$ in (26) and (27) with (23), we derive

$$\begin{aligned}
& \frac{ES_2}{ES_1}\left(\frac{\gamma_2 \hat{h}_1+\theta_2}{\gamma_1 \hat{h}_1+\theta_1}\right)^{-1}\left(\frac{q_2}{q_1}\right)^{\eta-1} \\
= \; & \frac{(1-\alpha_2)}{(1-\alpha_1)} \frac{\lambda_2\left(\gamma_2 \exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_2\right)}{\lambda_1\left(\gamma_1 \exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_1\right)},
\end{aligned} \tag{P2-5}$$

$$\begin{aligned}
& \frac{ES_2}{ES_3}\left(\frac{\gamma_2 \hat{h}_2+\theta_2}{\gamma_3 \hat{h}_2+\theta_3}\right)^{-1}\left(\frac{q_2}{q_3}\right)^{\eta-1} \\
= \; & \frac{(1-\alpha_2)}{(1-\alpha_3)} \frac{\lambda_2\left(\gamma_2 \exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_2\right)}{\lambda_3\left(\gamma_3 \exp\left(\mu+\frac{1}{2}\sigma^2\right)+\theta_3\right)},
\end{aligned} \tag{P2-6}$$

where the right-hand sides of both equations are constant.

Given that (P2-1)-(P2-6) must hold, what remains of the proof is to show that it must be the case that

$$
\begin{aligned}
\frac{\partial \hat{h}_1}{\partial \xi_x} &> 0, \\
\frac{\partial \hat{h}_2}{\partial \xi_x} &< 0.
\end{aligned}
$$

In addition, the following properties will be repeatedly used in the proof:

$$
\frac{\partial ES_1}{\partial \hat{h}_1} > 0, \quad \frac{\partial ES_2}{\partial \hat{h}_1} < 0, \quad \frac{\partial ES_2}{\partial \hat{h}_2} > 0, \quad \frac{\partial ES_3}{\partial \hat{h}_2} < 0; \tag{30}
$$

$$
\begin{aligned}
\frac{\partial \left(\frac{\gamma_2 \hat{h}_1 + \theta_2}{\gamma_1 \hat{h}_1 + \theta_1}\right)}{\partial \hat{h}_1} &= \gamma_2\theta_1 - \gamma_1\theta_2 = \left(\varepsilon_1^\gamma + \varepsilon_1^\theta + \varepsilon_1^\gamma \varepsilon_1^\theta\right)\left(1 + \varepsilon_2^\theta\right)\gamma_1\theta_3 > 0, && (31) \\
\frac{\partial \left(\frac{\gamma_2 \hat{h}_2 + \theta_2}{\gamma_3 \hat{h}_2 + \theta_3}\right)}{\partial \hat{h}_2} &= \gamma_2\theta_3 - \gamma_3\theta_2 = -\left(\varepsilon_2^\gamma + \varepsilon_2^\theta + \varepsilon_2^\gamma \varepsilon_2^\theta\right)\left(1 + \varepsilon_1^\gamma\right)\gamma_1\theta_3 < 0; && (32)
\end{aligned}
$$

$$
\begin{aligned}
ES_1 &= \Phi\left(\frac{\ln \hat{h}_1 - \mu}{\sqrt{2}\sigma}\right), && (33) \\
ES_2 &= \Phi\left(\frac{\ln \hat{h}_2 - \mu}{\sqrt{2}\sigma}\right) - \Phi\left(\frac{\ln \hat{h}_1 - \mu}{\sqrt{2}\sigma}\right), && (34) \\
ES_3 &= 1 - \Phi\left(\frac{\ln \hat{h}_2 - \mu}{\sqrt{2}\sigma}\right), && (35)
\end{aligned}
$$

where Φ denotes the CDF of the standard normal distribution. We then prove by contradiction.

1. Suppose that the following conditions hold:

$$
\frac{\partial \hat{h}_1}{\partial \xi_x} < 0 \text{ and } \frac{\partial \hat{h}_2}{\partial \xi_x} > 0. \tag{36}
$$

(36) immediately implies the following:

$$\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right) > 0, \quad \frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_3}\right) > 0;$$
$$\frac{\partial}{\partial \xi_x}\left(\frac{\gamma_2 \hat{h}_1 + \theta_2}{\gamma_1 \hat{h}_1 + \theta_1}\right) < 0, \quad \frac{\partial}{\partial \xi_x}\left(\frac{\gamma_2 \hat{h}_2 + \theta_2}{\gamma_3 \hat{h}_2 + \theta_3}\right) < 0.$$

Thus, combined with (P2-5) and (P2-6), it implies that an increase in ξ_x results in decreases in $\left(\frac{q_2}{q_1}\right)^{\eta-1}$ and $\left(\frac{q_2}{q_3}\right)^{\eta-1}$. Since $\eta < 1$, we thus have

$$\frac{\partial}{\partial \xi_x}\left(\frac{q_2}{q_1}\right) > 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{q_2}{q_3}\right) > 0.$$

If $\frac{q_2}{q_1}$ increases, $\left(q_2^{\eta+\frac{\alpha_2}{1-\alpha_2}} / q_1^{\eta+\frac{\alpha_1}{1-\alpha_1}}\right) = \left(\frac{q_2}{q_1}\right)^{\eta}\left(q_2^{\frac{\alpha_2}{1-\alpha_2}} / q_1^{\frac{\alpha_1}{1-\alpha_1}}\right)$ also increases, as $\eta > 0$ and $\frac{\alpha_2}{1-\alpha_2} > \frac{\alpha_1}{1-\alpha_1}$; for $\left(q_2^{\eta+\frac{\alpha_2}{1-\alpha_2}} / q_3^{\eta+\frac{\alpha_3}{1-\alpha_3}}\right)$, the argument is similar. This contradicts (24) and (25) because we have shown that both $\frac{ES_2}{ES_1}$ and $\frac{ES_2}{ES_2}$ increase. Therefore, it cannot be the case that $\frac{\partial \hat{h}_1}{\partial \xi_x} < 0$ and $\frac{\partial \hat{h}_2}{\partial \xi_x} > 0$.

2. Suppose that the following conditions hold:

$$\frac{\partial \hat{h}_1}{\partial \xi_x} < 0 \text{ and } \frac{\partial \hat{h}_2}{\partial \xi_x} < 0. \tag{37}$$

We examine the all the possible cases and show that none can hold.

(a) Suppose that

$$\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right) > 0.$$

Applying the assumptions that $\frac{\partial \hat{h}_1}{\partial \xi_x} < 0$ and $\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right) > 0$ to (31) and (P2-5) results in $\frac{\partial}{\partial \xi_x}\left(\frac{q_2}{q_1}\right) > 0$ because $\eta < 1$. We know that $\left(q_2^{\eta+\frac{\alpha_2}{1-\alpha_2}} / q_1^{\eta+\frac{\alpha_1}{1-\alpha_1}}\right)$ increases if $\frac{q_2}{q_1}$ increases, as explained earlier in part 1. This thus results in a contradiction of (24).

(b) Suppose that

$$\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right) < 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_3}\right) > 0.$$

Combining (33)-(34) and the assumption $\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right) < 0$, we have

$$\frac{\frac{\partial}{\partial \xi_x} ES_2}{ES_2} - \frac{\frac{\partial}{\partial \xi_x}\Phi\left(\frac{\ln \hat{h}_1 - \mu}{\sqrt{2}\sigma}\right)}{ES_1} < 0.$$

We know that $\frac{\partial}{\partial \xi_x}\Phi\left(\frac{\ln \hat{h}_1 - \mu}{\sqrt{2}\sigma}\right) < 0$ given $\frac{\partial \hat{h}_1}{\partial \xi_x} < 0$, and accordingly, we have

$$\frac{\frac{\partial}{\partial \xi_x} ES_2}{ES_2} < 0.$$

Moreover, we derive

$$\frac{\partial}{\partial \xi_x}\left(\log \frac{ES_2}{ES_3}\right) = \frac{\frac{\partial}{\partial \xi_x} ES_2}{ES_2} + \frac{\frac{\partial}{\partial \xi_x}\Phi\left(\frac{\ln \hat{h}_2 - \mu}{\sqrt{2}\sigma}\right)}{ES_3},$$

and the first argument is thus negative; the second argument is also negative given $\frac{\partial \hat{h}_2}{\partial \xi_x} < 0$. Therefore, we have $\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_3}\right) < 0$. This results in a contradiction.

(c) Suppose that

$$\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right) < 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_3}\right) < 0.$$

Taking the log of (P2-5) and then taking derivative with respect to ξ_x

yields

$$\frac{\frac{\partial\Phi\left(\frac{\ln\hat{h}_2-\mu}{\sqrt{2}\sigma}\right)}{\partial\hat{h}_2}\frac{\partial\hat{h}_2}{\partial\xi_x}-\frac{\partial\Phi\left(\frac{\ln\hat{h}_1-\mu}{\sqrt{2}\sigma}\right)}{\partial\hat{h}_1}\frac{\partial\hat{h}_1}{\partial\xi_x}}{\Phi\left(\frac{\ln\hat{h}_2-\mu}{\sqrt{2}\sigma}\right)-\Phi\left(\frac{\ln\hat{h}_1-\mu}{\sqrt{2}\sigma}\right)}-\frac{\frac{\partial\Phi\left(\frac{\ln\hat{h}_1-\mu}{\sqrt{2}\sigma}\right)}{\partial\hat{h}_1}\frac{\partial\hat{h}_1}{\partial\xi_x}}{\Phi\left(\frac{\ln\hat{h}_1-\mu}{\sqrt{2}\sigma}\right)}-\frac{G_1\left(\varepsilon^\gamma,\varepsilon^\theta\right)\frac{\partial\hat{h}_1}{\partial\xi_x}}{\left(\frac{\gamma_2\hat{h}_2+\theta_2}{\gamma_3\hat{h}_2+\theta_3}\right)}$$
$$-\left(1-\eta\right)\frac{\frac{\partial}{\partial\xi_x}\left(\frac{q_2}{q_1}\right)}{\left(\frac{q_2}{q_1}\right)}=0,$$

where we have applied (31), and define

$$G_1\left(\varepsilon^\gamma,\varepsilon^\theta\right)=\left(\varepsilon_1^\gamma+\varepsilon_1^\theta+\varepsilon_1^\gamma\varepsilon_1^\theta\right)\left(1+\varepsilon_2^\theta\right)\gamma_1\theta_3.$$

After rearrangement, we derive

$$\left(1-\eta\right)\frac{\frac{\partial}{\partial\xi_x}\left(\frac{q_2}{q_1}\right)}{\left(\frac{q_2}{q_1}\right)}=-\left(\frac{G_1\left(\varepsilon^\gamma,\varepsilon^\theta\right)}{\left(\frac{\gamma_2\hat{h}_2+\theta_2}{\gamma_3\hat{h}_2+\theta_3}\right)}+\frac{\frac{\partial\Phi\left(\frac{\ln\hat{h}_1-\mu}{\sqrt{2}\sigma}\right)}{\partial\hat{h}_1}}{ES_2}+\frac{\frac{\partial\Phi\left(\frac{\ln\hat{h}_1-\mu}{\sqrt{2}\sigma}\right)}{\partial\hat{h}_1}}{ES_1}\right)\frac{\partial\hat{h}_1}{\partial\xi_x}$$
$$+\frac{\frac{\partial\Phi\left(\frac{\ln\hat{h}_2-\mu}{\sqrt{2}\sigma}\right)}{\partial\hat{h}_2}}{ES_2}\frac{\partial\hat{h}_2}{\partial\xi_x}.$$

Since $\frac{\partial\hat{h}_1}{\partial\xi_x}<0$ and $\eta<1$, there exists an $\hat{\varepsilon}$ such that $\frac{\partial}{\partial\xi_x}\left(\frac{q_2}{q_1}\right)>0$ if

$$\varepsilon_j^i>\hat{\varepsilon}\quad\text{for some }\ i\in\{\gamma,\theta\},\ j\in\{1,2\}.\tag{38}$$

Moreover, we have $\frac{\partial}{\partial\xi_x}\left(q_2^{\eta+\frac{\alpha_2}{1-\alpha_2}}/q_1^{\eta+\frac{\alpha_1}{1-\alpha_1}}\right)>0$ since $\eta>0$ and $\frac{\alpha_2}{1-\alpha_2}>\frac{\alpha_1}{1-\alpha_1}$, as explained in part 1. In this case, let $\tilde{Q}_{21}\left(\xi_x\right)$ denote $\left(q_2^{\eta+\frac{\alpha_2}{1-\alpha_2}}/q_1^{\eta+\frac{\alpha_1}{1-\alpha_1}}\right)$, and we can express it as follows:

$$\frac{\partial}{\partial\xi_x}\tilde{Q}_{21}\left(\xi_x\right)=\tilde{G}_1\left(\varepsilon^\gamma,\varepsilon^\theta\right),$$

where $\tilde{G}_1\left(\varepsilon^{\gamma},\varepsilon^{\theta}\right)$ is increasing in $\left(\varepsilon_j^{\gamma},\varepsilon_j^{\theta}\right)$:

$$\frac{\partial \tilde{G}_1\left(\varepsilon^{\gamma},\varepsilon^{\theta}\right)}{\partial \varepsilon_j^{\gamma}} \geq 0, \quad \frac{\partial \tilde{G}_1\left(\varepsilon^{\gamma},\varepsilon^{\theta}\right)}{\partial \varepsilon_j^{\theta}} \geq 0 \quad \text{for} \quad j \in \{1,2\}. \tag{39}$$

Taking the log of $\frac{ES_2}{ES_1}\tilde{Q}_{21}\left(\xi_x\right)$ and then taking derivative with respect to ξ_x yields

$$\frac{\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right)}{\frac{ES_2}{ES_1}} + \frac{\tilde{G}_1\left(\varepsilon^{\gamma},\varepsilon^{\theta}\right)}{\tilde{Q}_{21}\left(\xi_x\right)}. \tag{40}$$

According to (39), we know that there exists an $\hat{\varepsilon}'$ such that (40)>0 if $\varepsilon_j^i > \hat{\varepsilon}'$ for some $i \in \{\gamma,\theta\}$, $j \in \{1,2\}$. This results in a contradiction of (24).

3. Suppose that the following conditions hold:

$$\frac{\partial \hat{h}_1}{\partial \xi_x} < 0 \text{ and } \frac{\partial \hat{h}_2}{\partial \xi_x} < 0. \tag{41}$$

The analysis is symmetric to part 2 and hence omitted.

We have shown that no cases specified in parts 1 to 3 can hold. By using proof by contradiction, we have proved that it must be the case that $\frac{\partial \hat{h}_1}{\partial \xi_x} > 0$ and $\frac{\partial \hat{h}_2}{\partial \xi_x} < 0$. *QED*

B.3 Proof of Corollary 1

From Proposition B.2, we have proven that $\frac{\partial \hat{h}_1}{\partial \xi_x} > 0$ and $\frac{\partial \hat{h}_2}{\partial \xi_x} < 0$, which, combined with (30)-(32), immediately yield the following:

$$\begin{aligned} \frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right) &< 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_3}\right) < 0, \\ \frac{\partial}{\partial \xi_x}\left(\frac{\gamma_2\hat{h}_1+\theta_2}{\gamma_1\hat{h}_1+\theta_1}\right) &> 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{\gamma_2\hat{h}_2+\theta_2}{\gamma_3\hat{h}_2+\theta_3}\right) > 0. \end{aligned}$$

Combining these equations with (P2-5)-(P2-6), and since $\eta < 1$, we obtain

$$\frac{\partial}{\partial \xi_x}\left(\frac{q_2}{q_1}\right) < 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{q_2}{q_3}\right) < 0. \tag{C1-1}$$

Combining (C1-1) with (14), and since $\eta > 0$, we further obtain

$$\frac{\partial}{\partial \xi_x}\left(\frac{Y_2}{Y_1}\right) > 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{Y_2}{Y_3}\right) > 0. \tag{C1-2}$$

Finally, from (19), (5), and the definitions that $K_j = \int_{\mathbb{O}_j} k_j\,(h)\,dF(h)$, $Y_j = \int_{\mathbb{O}_j} y_j\,(h)\,dF(h)$, we derive the relationship between K_j and Y_j:

$$\frac{K_j}{K_{j'}} = \frac{Y_j}{Y_{j'}}\frac{q_j}{q_{j'}}\frac{\alpha_j}{\alpha_{j'}},$$

and by combining it with (14), we obtain

$$\frac{K_j}{K_{j'}} = \left(\frac{q_j}{q_{j'}}\right)^{1-\eta}\frac{\alpha_j}{\alpha_{j'}}\left(\frac{\lambda_j}{\lambda_{j'}}\right)^{-1}.$$

Since $\eta < 1$, and according to (C1-1), we obtain

$$\frac{\partial}{\partial \xi_x}\left(\frac{K_2}{K_1}\right) < 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{K_2}{K_3}\right) < 0. \tag{C1-3}$$

QED

B.4 Proof of Proposition 3 (Wage Polarization)

Given (12) and (13), the average wages specified in (15) can be expressed as

$$\overline{w}_j = q_j^{\frac{1}{1-\alpha_j}}\xi_x^{\frac{\alpha_j}{1-\alpha_j}}(1-\alpha_j)\left(\frac{\alpha_j}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}}\left(\gamma_j \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_j\right). \tag{42}$$

Combining this with (26) and (27), we derive the average wage of middle-skill workers relative to that of low- and high-skill workers as follows:

$$\begin{aligned}
\frac{\overline{w}_2}{\overline{w}_1} &= \left(\frac{\gamma_1\hat{h}_1 + \theta_1}{\gamma_2\hat{h}_1 + \theta_2}\right)\frac{\gamma_2 \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_2}{\gamma_1 \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_1},\\
\frac{\overline{w}_2}{\overline{w}_3} &= \left(\frac{\gamma_3\hat{h}_2 + \theta_3}{\gamma_2\hat{h}_2 + \theta_2}\right)\frac{\gamma_2 \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_2}{\gamma_3 \exp\left(\mu + \frac{1}{2}\sigma^2\right) + \theta_3}.
\end{aligned}$$

The proof of Proposition 2 demonstrates that $\frac{\partial \hat{h}_1}{\partial \xi_x} > 0$ and $\frac{\partial \hat{h}_2}{\partial \xi_x} < 0$ if $\eta < 1$. Accordingly, we derive the following:

$$\begin{aligned}
\frac{\partial\left(\frac{\overline{w}_2}{\overline{w}_1}\right)}{\partial \xi_x} &= \frac{\partial\left(\frac{\overline{w}_2}{\overline{w}_1}\right)}{\partial\left(\frac{\gamma_1\hat{h}_1+\theta_1}{\gamma_2\hat{h}_1+\theta_2}\right)} \frac{\partial\left(\frac{\gamma_1\hat{h}_1+\theta_1}{\gamma_2\hat{h}_1+\theta_2}\right)}{\partial \hat{h}_1} \frac{\partial \hat{h}_1}{\partial \xi_x} < 0, \\
\frac{\partial\left(\frac{\overline{w}_2}{\overline{w}_3}\right)}{\partial \xi_x} &= \frac{\partial\left(\frac{\overline{w}_2}{\overline{w}_3}\right)}{\partial\left(\frac{\gamma_3\hat{h}_2+\theta_3}{\gamma_2\hat{h}_2+\theta_2}\right)} \frac{\partial\left(\frac{\gamma_3\hat{h}_2+\theta_3}{\gamma_2\hat{h}_2+\theta_2}\right)}{\partial \hat{h}_2} \frac{\partial \hat{h}_2}{\partial \xi_x} < 0,
\end{aligned}$$

where we have applied (31) and (32). *QED*

B.5 Proof of Proposition 4 (Task-specific Productivity)

For $j' = 1, 3$, we derive

$$\begin{aligned}
\log \frac{\hat{y}_2}{\hat{y}_{j'}} &= \log \frac{Y_2}{Y_{j'}} - \log \frac{ES_2}{ES_{j'}} \\
&= -\eta \log \left(\frac{q_2}{q_{j'}}\right) - \log \frac{ES_2}{ES_{j'}},
\end{aligned}$$

where the second equality is derived by applying (14). To prove the proposition, we need to show that $\frac{\partial}{\partial \xi_x}\left(\log \frac{\hat{y}_2}{\hat{y}_{j'}}\right) < 0$. According to Proposition 2, we know that $\frac{\partial \hat{h}_1}{\partial \xi_x} > 0$ and $\frac{\partial \hat{h}_2}{\partial \xi_x} < 0$ must hold. Combining the two properties with (P2-5)-(P2-6), (30)-(32), and given $\eta < 1$, we thus have the following:

$$\frac{\partial}{\partial \xi_x}\left(\frac{q_2}{q_1}\right) < 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{q_2}{q_3}\right) < 0;$$

$$\frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_1}\right) < 0 \text{ and } \frac{\partial}{\partial \xi_x}\left(\frac{ES_2}{ES_3}\right) < 0.$$

Therefore, we derive that for $j = 1, 3$,

$$\frac{\partial}{\partial \xi_x}\left(\log \frac{\hat{y}_2}{\hat{y}_{j'}}\right) > 0$$

since $\eta > 0$. *QED*

Appendix C Quantitative

C.1 Task-specific Moments

The labor share of task j is defined as

$$LS_j \equiv \frac{\int_{\mathbb{O}_j} w_j(h)\, dF(h)}{q_j \int_{\mathbb{O}_j} y_j(h)\, dF(h)}. \tag{43}$$

From (12), the total wages in task j are derived as

$$\int_{\mathbb{O}_j} w_j(h)\, dF(h) = q_j^{\frac{1}{1-\alpha_j}} (1-\alpha_j) \left(\frac{\alpha_j \xi_x}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}} \int_{\mathbb{O}_j} \left(\gamma_j h + \theta_j\right) dF(h).$$

From (21), the total output in task j is derived as

$$\int_{\mathbb{O}_j} y_j(h)\, dF(h) = q_j^{\frac{\alpha_j}{1-\alpha_j}} \left(\frac{\alpha_j \xi_x}{r+\delta}\right)^{\frac{\alpha_j}{1-\alpha_j}} \int_{\mathbb{O}_j} \left(\gamma_j h + \theta_j\right) dF(h).$$

Given the above, we derive that the labor share of task j defined in (43) equals $1-\alpha_j$.

Let I_j denote the income share of task j, and it is defined as follows:

$$I_j = \frac{\overline{w}_j ES_j}{\sum_{j \in J} \overline{w}_j ES_j} \tag{44}$$

where $\overline{w}_j$ the average wage of workers in task j, and ES_j is the employment share of task j.

C.2 Calibration

Table A.1: Calibrated Parameters

		Alternative (NIPA)
Note	Parameter	Value
Task production	γ_2	5.42
	γ_3	57.35
	θ_1	272.32
	θ_2	8.23
Good production	η	0.0163
Skill distribution	σ	1.966

The table reports the calibrated parameter values using the relative price of the investment good based on the NIPA deflator.

C.3 Quantitative Results

Table A.2: Model Fit: Alternative

Variable	Data 1980	Data 2010	Data Δ of var.	Model 1980	Model 2010	Model Δ of var.	Model Fit % explained
ES_3	25.83	36.95	11.12	25.82	29.65	3.83	34.44%
ES_2	61.35	45.89	-15.46	61.34	53.97	-7.37	47.67%
ES_1	12.83	17.16	4.33	12.84	16.39	3.55	81.99%
$\frac{\overline{w}_2}{\overline{w}_3}$	0.6519	0.4745	-27.2%	0.6518	0.5533	-15.1%	55.51%
$\frac{\overline{w}_2}{\overline{w}_1}$	1.6362	1.3720	-16.1%	1.6362	1.6303	-0.36%	2.24%
I_3	36.41	57.14	20.73	36.41	45.56	9.15	44.14%
I_2	56.38	33.68	-22.70	56.38	45.89	-10.49	46.21%
I_1	7.21	9.18	1.97	7.21	8.55	1.34	68.02%
Labor share	0.568	0.533	-6.16%	0.633	0.615	-2.84%	

The table reports the results using the relative price of the investment good based on the NIPA deflator.

Table A.3: Decomposition: Alternative

Variable	α_j invariant % explained	R invariant % explained	Decomposition R contributes	Decomposition α_j contributes
ES_3	9.8%	18.97%	41%	59%
ES_2	11.64%	25.94%	40%	60%
ES_1	16.17%	43.88%	39%	61%
$\frac{\overline{w}_2}{\overline{w}_3}$	16.92%	31.93%	39%	61%
$\frac{\overline{w}_2}{\overline{w}_1}$	0.42%	1.18%	39%	61%
I_3	12.30%	24.22%	41%	59%
I_2	12.29%	25.42%	40%	60%
I_1	12.69%	38.07%	38%	62%

The table reports the quantitative results using the relative price of the investment good based on the NIPA deflator.

Appendix D Data

D.1 Industry-Occupation Matrix

Table A.4: High-skill and Low-skill Occupations

Occupation c	Management		Professional		Service	
	1980	2010	1980	2010	1980	2010
τ_c	10.237	14.495	15.588	22.454	12.827	17.161
Industry d	$\tau_{(c,d)}$		$\tau_{(c,d)}$		$\tau_{(c,d)}$	
Mining	0.99	0.64	0.75	0.26	0.15	0.03
Construction	4.31	8.56	1.69	0.74	0.40	0.24
Manufacturing	20.71	12.32	14.20	6.72	4.31	0.93
Wholesale trade	7.36	2.86	0.84	0.55	0.53	0.21
Retail trade	17.87	4.81	2.56	2.81	30.27	3.07
Transportation and utilities	4.98	3.83	1.71	1.10	2.04	1.29
Information	1.74	3.26	1.76	3.55	0.15	0.45
Financial activities	13.67	19.55	1.43	2.07	2.15	1.60
Professional and business services	9.79	17.55	10.77	16.69	4.59	13.34
Education and health services	9.94	14.92	55.95	59.72	31.26	31.54
Leisure and Hospitality	3.26	8.48	1.92	2.74	8.27	36.38
Other services	5.38	3.22	6.43	3.08	15.87	10.93

Data source: U.S. census.

Table A.5: Middle-skill Occupations

Occupation c	Sales		Production		Mechanics	
	1980	2010	1980	2010	1980	2010
τ_c	10.875	11.821	15.056	6.083	4.087	3.646
Industry d	$\tau_{(c,d)}$		$\tau_{(c,d)}$		$\tau_{(c,d)}$	
Mining	0.11	0.05	0.78	0.51	2.04	1.29
Construction	0.57	0.56	2.38	1.66	4.98	10.78
Manufacturing	7.09	4.11	77.91	70.19	25.48	14.00
Wholesale trade	10.77	8.92	2.89	1.69	6.09	2.88
Retail trade	59.53	56.76	5.20	6.39	16.98	12.79
Transportation and utilities	1.65	0.77	3.31	3.74	9.97	11.37
Information	0.57	2.63	0.21	0.82	8.69	6.43
Financial activities	14.31	13.54	0.32	0.70	0.50	2.96
Professional and business services	1.91	3.62	1.22	4.61	1.75	5.98
Education and health services	0.82	0.94	1.82	2.65	1.98	5.54
Leisure and Hospitality	1.10	5.82	0.45	1.41	0.90	2.73
Other services	1.57	2.29	3.52	5.64	20.64	23.26

Data source: U.S. census.

Table A.6: Middle-skill Occupations, continued

Occupation c	Construction		Administrative		Transportation	
	1980	2010	1980	2010	1980	2010
τ_c	5.725	5.432	16.984	12.746	8.619	6.163
Industry d	$\tau_{(c,d)}$		$\tau_{(c,d)}$		$\tau_{(c,d)}$	
Mining	6.26	3.82	0.72	0.34	2.55	1.16
Construction	71.32	82.12	2.48	2.79	6.21	2.78
Manufacturing	9.41	4.20	18.12	8.00	27.70	13.66
Wholesale trade	0.98	0.44	5.93	3.86	9.78	8.76
Retail trade	2.26	1.11	10.12	14.70	19.47	14.95
Transportation and utilities	4.87	2.75	9.84	9.89	26.03	37.75
Information	0.25	0.21	3.91	3.40	0.18	0.66
Financial activities	0.82	0.21	17.46	13.92	0.43	1.19
Professional and business services	0.76	2.33	8.71	13.43	1.20	6.60
Education and health services	1.87	1.49	16.98	22.28	3.20	5.50
Leisure and Hospitality	0.46	0.48	1.53	3.73	0.50	3.20
Other services	0.73	0.38	4.20	3.65	2.76	3.79

Data source: U.S. census.

REFERENCES

Acemoglu, D. and Autor, D. H. (2011). Skills, tasks and technologies: Implications for employment and earnings. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4B, pages 1043–1171. Elsevier.

Acemoglu, D. and Guerrieri, V. (2008). Capital deepening and nonbalanced economic growth. *Journal of Political Economy*, 116(3):467–498.

Autor, D. H. (2013). The "task approach" to labor markets: An overview. *Journal for Labour Market Research*, 46(3):185–199.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–1597.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2006). The polarization of the U.S. labor market. *American Economic Review*, 96(2):189–194.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in U.S. wage inequality: Revising the revisionists. *The Review of Economics and Statistics*, 90(2):300–323.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Bárány, Z. L. and Siegel, C. (2017). Job polarization and structural change. *American Economic Journal: Macroeconomics, Forthcoming*.

DeJong, D. N. and Dave, C. (2011). *Structural Macroeconometrics*. Princeton University Press, second edition.

Elsby, M. W. L., Hobijn, B., and Şahin, A. (2013). The decline of the U.S. labor share. *Brookings Papers on Economic Activity*.

Fisher, J. D. M. (2006). The dynamic effects of neutral and investment-specific technology shocks. *Journal of Political Economy*, 114(3):413–451.

Goos, M. and Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *The Review of Economics and Statistics*, 89(1):118–133.

Goos, M., Manning, A., and Salomons, A. (2009). Job polarization in Europe. *American Economic Review*, 99(2):58–63.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, 104(8):2509–2526.

Grossman, G. M., Helpman, E., Oberfield, E., and Sampson, T. (2017). Balanced growth despite Uzawa. *American Economic Review*, 107(4):1293–1312.

Herrendorf, B. and Schoellman, T. (2017). Wages, human capital, and barriers to structural transformation. *American Economic Journal: Macroeconomics, Forthcoming*.

Jones, C. I. (2005). The shape of production functions and the direction of technical change. *The Quarterly Journal of Economics*, 120(2):517–549.

Justiniano, A., Primiceri, G. E., and Tambalotti, A. (2009). Investment shocks and the relative price of investment. *Review of Economic Dynamics*, 14(1):102–121.

Karabarbounis, L. and Neiman, B. (2014). The global decline of the labor share. *The Quarterly Journal of Economics*, 129(1):61–103.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V., and Violante, G. L. (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, 68(5):1029–1053.

Lawrence, R. Z. (2015). Recent declines in labor's share in US income: A preliminary neoclassical account. *NBER Working Paper*, (No. 21296).

León-Ledesma, M. A. and Satchi, M. (2017). Appropriate technology and balanced growth. *The Review of Economic Studies, Forthcoming*.

Ngai, L. R. and Pissarides, C. A. (2007). Structural change in a multisector model of growth. *American Economic Review*, 97(1):429–443.

Stokey, N. L. (2016). Technology, skill and the wage structure. *NBER Working Paper*, (No. 22176).

Uras, B. R. and Wang, P. (2017). Production flexibility, misallocation and total factor productivity. *NBER Working Paper*, (No. 23970).