Does Digitalization Widen Labor Income Inequality?*

Jilei Huang †
Shandong University

Yi Sun‡
University of Chinese Academy of Sciences

Liyan Yang§
University of Toronto

December, 2023

Abstract

Many studies have suggested a positive and monotonic relationship between technological progress and wage income inequality since 1980s for industrialized economies due to the well-known replacement effect on labor demand. We re-examine this issue by noting that the new generation of information technology, i.e., digitalization, can bring about offsetting labor supply effects. We demonstrate this idea in a growth model with human capital where digitalization endogenously shifts the worker ability distribution as a creative destruction effect and reduces the costs of learning. These two effects affect labor supply and occupational choice and reduce wage income gap between groups thereby countervail the traditional labor demand effect. As a result, the pattern of inequality-digitalization relationship can show nonlinear dynamics including an inverted U-shape. Using a recent panel dataset of the Chinese economy where digitalization has gained prominent growth, we conduct empirical test of our hypotheses and find supportive evidence of a "digital Kuznets curve". Our study contributes to the understanding of the nature of digitalization in re-shaping labor market structure.

Keywords: digitalization, wage income inequality, human capital, worker ability, learning cost

JEL Reference Numbers: O33, O34, O15, E17, O41, O47, J01

---

*We thank the valuable comments from discussants and participants in the Asian Meeting of the Econometric Society, China Financial Research Conference, China International Conference in Finance. The authors thank for the financial support of the Seeds Funds for International Research Collaboration by Shandong University. All views and errors are ours.

†Address: School of Economics, Shandong University, Jinan, 250100; e-mail: jilei.huang@sdu.edu.cn

‡Address: School of Economics and Management, University of Chinese Academy of Sciences, Beijing, China; e-mail: suny@ucas.ac.cn

§Address: Rotman School of Management, University of Toronto, Toronto, Canada; e-mail: liyan.yang@rotman.utoronto.ca

1 Introduction

With the recent waves of new technology development such as artificial intelligence (AI), robotics and big data, the world economy is experiencing an on-going digital transformation. A central question about this digitalization is what labor market consequences it can bring about. Looking back to the literature on the labor market implications of past generations of technologies, researchers have achieved large consensus that new technologies enlarges wage income gap between groups of workers since 1980s (see, e.g., Goldin and Katz (2009), Acemoglu and Autor (2011), Piketty (2014), Acemoglu Restrepo (2022))[1]. Does this prediction remain valid for the digital era? This paper examines the impact of digitalization on wage income inequality by building a growth model with digitalization and a labor market with heterogenous human capital.

To build up our analysis, we note first that much of the aforementioned work has relied on the labor demand side effect of technologies where the replacement effect always dominates. Whilst this could be largely true for technologies developed from the first to third industrial revolutions, the new generation of technology may exert significant labor supply effects. For example, technologies affect also occupational choice and labor mobility of workers (Galor and Moav 2000, Kambourov and Manovskii 2009, Cortes and Gallipoli 2018). For another, newly emerged technologies spread out knowledge of production and helps less skilled workers learn faster (Aghion 2002, He and Liu 2008, Restuccia and Vandenbroucke 2013, Brynjolfsson *et. al.* 2023).

We develop a unified theory to understand the mechanism underlying the relationship between digitalization and wage income inequality by incorporating both labor demand and labor supply side effects of digitalization. A task-based growth model with heterogenous human capital is then built to illustrate the main idea. In this model, digitalization works in a similar way with automation and its effects on capital and labor depends on the elasticity of substitution between the two factors (See, e.g., Aghion *et al.* (2017)). We consider the case that digitalization is capital depleting in the long-run and employ a capital-augmenting technology to achieve a balanced growth path (See, e.g., Grossman *et al.* (2017)). For labor market, workers abilities follow a uniform distribution and they accumulate human capital and make occupational choice as in Galor and Moav (2000). Digitalization then affects the labor market structure through three channels. One is the labor demand replacement effect caused by the skill-biased technological change (SBTC,

[1] Influences of new technologies on other aspects of labor market, such as labor share and capital income are much more controversial. For example, while replacements of labor by new technologies are confirmed in several studies (e.g., Acemoglu and Restrepo 2020), some positive effects of new technologies remain valid (see, e.g., Acemoglu and Restrepo 2020, and Hemous and Olsen 2022).

e.g., Katz and Murphy 1992, Acemoglu 1998). The second channel is that worker ability distribution is interrupted by digitalization, a creative destruction effect on labor supply. The third channel is that digitalization reduces the costs of learning of workers who choose to become skilled ones. Importantly, the latter two channels are novel and they affect occupational choice and wage income of two groups of workers endogenously.

We formulate several propositions to illustrate how these two new features contribute to the dynamic impact of digitalization on between-group and within-group wage income inequalities. Specifically, we show that, when digitalization also shrinks the worker ability distribution, it reduces between-group wage income inequality. This is also true when digitalization helps reduce learning costs of workers. The creative destruction channel and learning cost channel differ in their impacts on within-group inequalities, i.e., the former mechanism decreases (increases) within-group inequality of skilled (unskilled) workers, whilst the latter mechanism increases (decreases) within-group inequality of skilled (unskilled) workers. When putting together both labor demand two labor supply channels, the dynamic relationship of wage income gap and digitalization depends on the relative strengths of each channel. In particular, the between-group wage income inequality can deliver an inverted U-shape when either (or both) labor supply channels are present and just offset the traditional labor demand effect. These theoretical predictions are then verified with numerical simulations in a later section.

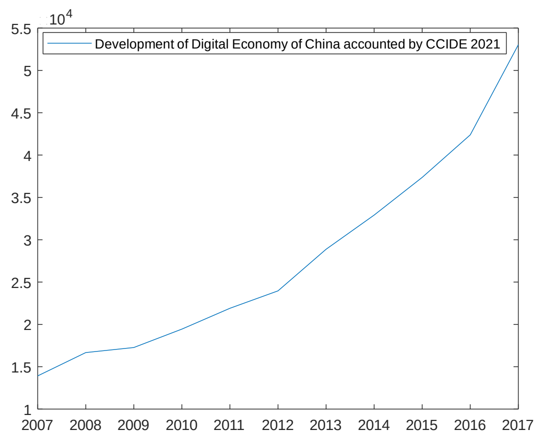

Figure 1: China's Digital Economy: 2007-2017 (100 million RMB)

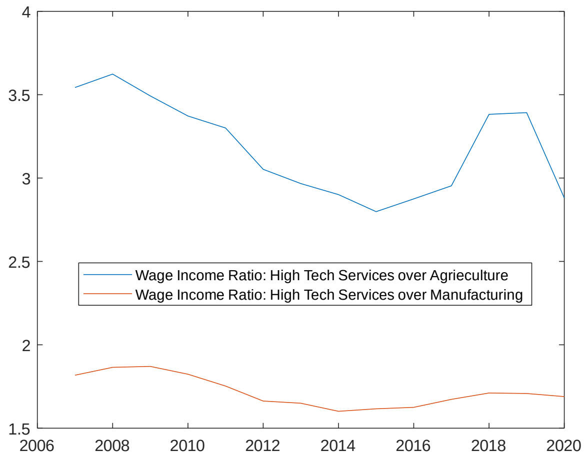

Figure 2: China's labor income inequalities

To see whether our theoretical model is a good description of the reality, we empirically test the relationship between digitalization and wage income inequality. The Chinese economy offers an exceptional opportunity to examine this issue. Despite being a follower of digital technologies and an emerging market, has achieved prominent progress in digitalization in the last decade in terms of digital transformation, automation and even robots production (see, e.g., Cheng *et al.* 2019)[2]. Figure 1 shows the ten-year average of the digital economy of China at aggregate level according to the "2021 Categorization of Core Industries of Digital Economy" (CCIDE 20212[3], thereafter) published by the National Bureau of Statistics of China (NBSC) in 2021. On the other hand, we do not observe a sharp increase in labor income gap as shown in Figure 2 where the sectoral wage gaps are plotted. This is in stark contrast to the widening wage gap observed in western countries. This calls for the need to re-examine the relationship between technological advancement and income inequality, both theoretically and empirically.

Using a constructed panel dataset, we conduct panel data regressions with two-way fixed effects and instrumental variables to quantify the impacts of digitalization on labor income inequality. The former is measured by the proportion of value-added of digital economy to local output, and the latter is measured by the wage income ratio of high-tech service sector over agricultural sector for the time period 2007-2019. In addition, to help understand the relationship, we also conducted mediating effect test to understand

[2]The Chinese government has also provided continuous supportive policies for digitalization. Examples of such policies include the Robotics Industry Development Plan (2016–2020) by Ministry of Industry and Information Technology (MIIT), “Made in China 2025” program and the fourteenth “5-Year Plan of Digital Economy Development” by the State Council.

[3]The CCIDE 2021 categorization has used the categorization of U.S. Bureau of Economic Analysis (BEA) and OECD for references.

the transmission mechanism of digitalization. The empirical results show that, first, there exists a Kuznets curve in Chinese digital economy which characterizes an inverted U-shaped relationship between digitalization and labor income inequality. The inflection point is around 8-10% of digitalization. This is in stark contrast with most of the empirical evidence of a monotonic positive relationship found in western countries.

The two labor supply channels we innovate are thus the key to understand the dynamic impacts of digitalization on labor income inequality. It has direct effect on wage income of both types of workers and indirect effect on occupational choice and labor supply. These effects are largely ignored in most research on technology-inequality in developed countries. This is not surprising given that empirical studies in the context of developed countries only found monotonic relationship. One exception is the experience of labor market of the US in the 1970s where between-group wage premium fell along with increases within-group wage gaps as a result of increases of labor supply caused by exogenous increases in college graduates. Our study then indicates that the Chinese digital economy in last decade may have experienced the same increase in labor supply as in 1970s for US, but it also has important difference in that it is not driven by increases in college students but a result of endogenous changes in labor ability and labor supply.

This paper contributes to the literature in three folds. First, it proposes labor supply effects of digitalization characterized by two novel mechanisms in re-shaping the dynamic relationship. The worker ability distribution channel which nests the positive effect of common technology and the negative effect of digitalization on worker ability. In particular, the creative destruction effect of digitalization shifts the distribution of worker ability to the left and thus becomes a countervailing force that offsets the positive monotonic impact of digitalization on inequality. The learning cost channel on the other hand shares spirit of positive externality of data economy. It diminishes wage income inequality through lowering the job threshold. These two mechanisms provide new insights to the nature and transmission channels of digitalization to macroeconomy. Second, it provides the first piece of empirical evidence on the nonlinear relationship between new technology and wage income inequality in the context of Chinese economy who is experiencing the most rapid digitalization in the world. The data we use is newly constructed in line with CCIDE published in 2021 and is consistent with national accounting of NBSC. Last but not least, our study points to dynamic understandings of digitalization and the importance of the supply behavior of labor whereas most existing literature only focus on labor demand. We thus deepen the understanding of the impacts of new technology on changes in labor structure in terms of human capital accumulation and labor choice. The findings of this study are helpful for understanding the nature of the impact of digitalization on labor market and have important implications for government policy.

The paper is organized as follows. Section 2 provides a literature review on the relationship between technology and inequality. Section 3 sets out a theoretical task-based growth model with human capital and endogenous occupational choice. Section 4 provides empirical evidence of the dynamic relationship between inequality and digitalization in the context of the Chinese economy. Section 5 conducts calibration and simulations of the theoretical model to demonstrate the dynamic relationships. Section 6 concludes.

2 Related Literature

This paper lies generally in the theme of examining the dynamic relationship between technology progress and inequality. The famous Kuznets Curve (Kuznets, 1955) depicts an inverted U-shaped relationship between economic growth and inequality. Since technology is the main driving force of economic growth, the Kuznets Curve implies also a hump-shaped relationship between technology and income inequality. However, this inverted U-Shape was soon overturned by researchers in 1990s due to the observation that Kuznets Curve only existed between 1915-1940s, after which the curve dropped substantially, remaining flat during 1960s-1970s, and sloped upward sharply since 1980s (Goldin and Katz 2009, Picketty 2014, Kasa and Lei (2018)). Therefore, in the long run, the relationship between technology and inequality looks more like an U-shape (e.g., also documented in Acemoglu and Autor 2011, Prettner and Schaefer 2021).

Researchers attribute the change in the shape of Kuznets curve to several factors, such as World War II (e.g., Milanovic, 2016), the rising of educated workers and higher human capital after 1940s (e.g., Goldin and Katz 2009), different intergenerational investments in education (Prettner and Schaefer 2021), and Skill-biased Technical Change (SBTC) after 1980s (e.g., Katz and Murphy 1992, Acemoglu 1998). Since 1980s, a bunch of 'new technologies' has broken through and diffused across the economy. Nonetheless, the technology-inequality relationship has remained increasing. Recent research paradigm on this issue largely follows the SBTC approach but taking new features of the new technologies into account. For example, technology may be still skill-biased in that human capital is more complementary with it (Galor and Moav 2000). For another, new technology can endogenously choose to augment capital or labor depending on its nature, following the Directed Technical Change (DCT) literature (Acemoglu 1998). Moreover, high-skilled labor may be complements with robots (Hemous and Olsen 2022). Among these new research directions, task-based model has gained popularity due to its flexibility and to that it can generate replacement effect that is absent in DCT models. Summarizing this strand of literature, we find little evidence of Kuznets curve after 1980s in developed countries.

Our study also sheds light on the recent debate on labor market consequences of digital technologies. In the early stage of information revolution, Kuhn and Mansour (2014) find empirical evidence that internet usage reduces job search costs. Regarding the recent wave of automation and artificial intelligence, Acemoglu and Restrepo (2018, 2020, 2022) study the impacts of adoptions of robots on US labor market using industrial and firm-level data. They find that labor share has fallen and the inequality between workers has enlarged as a result of automation. Those new trends of labor market outcomes are then explained in several task-based growth models which allow for various substitutional and compensational effects of automation. Frey (2020) compare different waves of automation and their different impacts on capital and labor. Korinek and Juelfs (2022) critically analyze the displacement effect of AI and automation and propose cases where human labor are still needed in future. Gomes *et al.* (2022) use an unique Sweden dataset which can capture people's exposure to robots, wealth rankings and demographics. They find quite large effects of automation on wealth distribution through a portfolio adjustment mechanism. We extend this literature by providing a new evidence in the context of the Chinese economy where continuous digitalization has been on going. Our theoretical framework, particularly the way we model digitalization in the production function share many similarities with that of modelling automation, but we allow for its direct and interactive effects on labor market.

In the aforementioned literature, the emphasis has been put on changes in labor demand. This is not surprising as technologies usually come out of activities of firms. However, labor supply also plays important roles in occupational choices and labor mobility. Examples of labor supply analysis include Katz and Murphy (1992), Acemoglu (1998), Aghion (2002), Grossman *et al.* (2017) and Hemous and Olsen (2022). Occupational choice is explicitly modelled in Galor and Moav (2000), Kambourov and Manovskii (2008, 2009), Dvorkin and Naranjo (2019). Costs of labor mobility has been discussed in Cortes and Gallipoli (2018). For developing countries, Du *et al.* (2014) and Ge and Yang (2014) document the continued influx of rural migrant workers to the industrial sector and their contribution to the productivity of labor-intensive industries. Therefore, our contribution to labor market literature by examining how digitalization affects workers' labor supply behavior.

The idea that digitalization affects human capital of workers is also directly related to the literature of the economics of data, especially in the externality and nonrivalry properties[4] and its consequences on growth, welfare and inequality. Jones et. al. (2020) examines

[4]Ghosh et. al. (2021) identifies the impact of digitalization on financial market. They demonstrate

the nonrivalry property of data and the positive externality it generates for society, and analyze its tradeoff with privacy in different ownership settings and government regulation schemes. Their finding is that authorize data ownership to consumers rather than firms delivers much higher welfare close to optimum as firms have less incentives to sell data to others. Cong et. al. (2021)[5] feed consumer generated data into production of goods to allow for semi-endogenous growth and find that on the balanced growth path a decentralized economy incurs welfare loss due to underemployment and overuse of data compared with the social optimum. Their model implies that *income gap* in the digital economy could be enlarged especially during the early stage of the balance growth path because economies starting from low initial growth generate low volume of data. In contrast, Ichihashi (2021), Acemoglu et. al. (2022) and Bergemann et. al. (2022) identify negative externalities, where the sharing of data of one consumer might reveal other consumers data, causing inefficient over-provision of data. Liu et. al. (2023) uncovers a new negative externality due to behavioral biases such as the existence of consumers with self-control problems to temptation goods. This offsets nonrivalry property of data and calls for consumer privacy protection. In addition, due to the existence of weak-willed consumers, digitalization raises total efficiency but at a price of widening welfare gap between strong- and weak-will consumers, which they named a *'algorithmic inequality'* problem. Farboodi and Veldkamp (2022) proposes a dynamic equilibrium model of the data economy taking data as an input of production and as a state variable with depreciation. They show that the long-run dynamics of data resembles decreasing returns to scale but it displays increasing returns to scale in the short-run. In particular, *a poverty trap* presents in the short-run between and large and small firms due the former taking advantage of increasing returns of data they generate during growth. In a related work, Farboodi et. al. (2022) develops a cross-sectional measure of data and documents a new fact of *data divergence* that, as large firms get larger, they attract much more data than other firms. Our work shares similar spirits with the above researchers but differs in that we also notice the possibility that digital technology may change the way workers learn and their ability distribution and occupational choice in a dynamic way. In addition, although some of the above papers discussed the inequality implications of digitalization, none of them focuses on labor market outcomes. In this sense, our work is closely related to a recent work by Brynjolfsson et. al. (2023)

both theoretically and empirically that FinTech lenders can achieve higher efficiency of screening borrowers when the latter use cashless payments and generate transferrable and verifiable information.

[5] Cong et. al. (2022) extend to vertical nonrivalry of data and a fully endogenous growth model where consumer-generated data can be shared with both production and innovation sectors. The innovation sector dominates production sector in the matter of long-run growth due to its advantage that, beside its dynamic nonrivalry, it "desensitizes" raw data into knowledge which avoids consumers' privacy concerns.

which demonstrates that newly emerged generative AI transmits the tacit knowledge of more skilled workers to low skilled workers and helps newer workers learn things faster. Our empirical evidence and theoretical analysis in this paper suggest that these dynamic learning effects on workers human capital may have taken effect well before the emergence of Generative AI.

Finally, it is noted that, our finding of a nonlinear (including an inverted U-shaped) relationship between technology and inequality is not the only case in literature. For developed countries, Borghans and Weel (2007) find that computer adoption causes Inverted U-shape of wage inequality when allowing workers decide endogenously whether and when to adopt computer basing on cost-benefit analysis. Their results are consistent with German wage structure during 1980s. Bohm *et al.* (2015) develop a two sector production model with immobile labour supply and directed technical change toward high-skilled human capital. They show that public policies that subsidize higher education costs for high-skilled workers raise inequality (in terms of wage rates, consumption and income) in the short run (3 decades), whereas they are beneficial for low-skilled workers in the long run. But their work is purely theoretical without empirical evidence. For developing countries, Che and Zhang (2017) uses higher education expansion 2003 as a natural experiment to examine its impact on tech adoption and TFP. They also empirically estimated the impact of education expansion on wage premium (using industry consensus data, 1995 and 2004). They show that college expansion cause increases of supply of high-skilled labor which drags down wage premium despite the increased demand for high-skills positions the same time. The whole dynamics show a humped shape (inverted U-shape): first increasing (1999-2002), then decreasing (2003-2007). Their work however, is subject to two drawbacks: i) it does not provide theoretical explanation of the underlying mechanism; ii) In their paper, education is purely exogenous, not a consequence of endogenous choice. Ge and Yang (2014) use Chinese urban household survey data and find that capital accumulation, skill-biased technological change, and rural-urban migration to be the major forces behind the evolving wage structure in urban China. Importantly, their model simulation reveals that high school wage premium can show inverted U-Shape 1992-2007 (peak in 2001). They argue that this is consistent with the dramatic labor migration from rural to urban areas during 1990s. Our paper complements these work by examining wage premium across industries in the digital era.

Besides, there are two additional empirical work in support of the existence of Kuznets curve. Messina and Silva (2019) find inverted U-Shape for Latin America wage inequality 1995-2015. They propose education attainment as the main reason for this nonlinearity as it reduces wages of college and high school graduates, and also of more experienced workers. However, they also find that two thirds of wage inequality decline is due to within group wage inequality, therefore reveals only limited role of labor supply effect emphasized by our paper. Castello-Climent and Domenech (2021) show that human capital inequality (as share of population with no schooling) has inverted U-Shape relationship with labor income inequality (146 countries, 1950-2010), but SBTC could have offset the effect of the fall in human capital inequality over time. We complement this literature by providing a theoretical framework to understand the mechanisms underlining these empirical patterns.

3 Theoretical Model

In this section, we set out a theoretical model to help understand the dynamic relationship between digitalization and wage income inequality. Some recent growth models in the literature has noticed multiples effects of digital technology, e.g., labor displacement effects versus productivity effects and non-rivalry of data, but none of them examined the specific nonlinearity they may exert. Importantly, they focus mainly on the firm side effects of digital technology and how it affects labor demand, but ignored how it may change labor supply of workers. We build a growth model where digitalization creates countervailing forces both in labor demand and labor supply. Specifically, we start from a canonical task-based growth model (see, e.g., Aghion *et al.* 2018) with CES production function where digitalized and un-digitalized tasks are complements so that digitalization is actually capital-depleting and labor-augmenting. This helps the model to achieve a unique balanced growth path in the presence of digitalization and thus allows for neat analysis of labor market dynamics without worrying about explosive paths. On the labor supply side, which is the main focus of our paper, we model labor market as consisting of a uniform distribution of workers with different abilities (see, e.g., Galor and Moav 2000). They make endogenous occupational choices of skilled and unskilled jobs given their individual ability and digitalization. On the labor demand side, we employ the skill-biased technological change (SBTC) and let digitalization cause firms to demand more skilled than unskilled workers. The SBTC setting would induce monotonic and increasing wage income gap between the two groups of workers. However, our labor supply side is enriched with two novel features to offset this monotonic positive relationship dynamically. One feature is that workers must pay learning costs if he/she intends to become a skilled worker, thus making the two jobs imperfect substitutes. Critically, due to positive externality of digital technology, the learning cost function is decreasing in digitalization. The second feature is that digitalization exerts a depreciating effect on the abilities of (both) workers, probably due to a creative destruction effect (the insight that not only unskilled workers but also skilled workers are replaced by recent development of digitalization seems to be supported by the evidence in Brynjolfsson et. al. 2023)[6]. We model this feature by antagonizing the distribution of worker abilities as a negative function of digitalization, a form similar with Acemoglu and Restrepo (2018). We will show below that the two new features on the labor supply side are the key to understand the dynamic relationship between digitalization and wage income inequality.

3.1 Economic Environment

There are representative firms and heterogenous workers, i.e., skilled and unskilled, in the economy. Goods market and factor markets are all perfectly competitive. Workers are endowed with abilities that follow a distribution in the domain $[A_t - 1, A_t]$. They make occupational choices between skilled and unskilled workers at time t-1. Workers who decide to become skilled ones invest in human capital by paying learning cost. In period t, both skilled and unskilled workers work for firms and receive wage income. In time t+1, both workers retire and spend all of their income for consumption and investment. For simplicity, worker save a constant fraction of net wage income in period t. Firms on the other hand produce intermediate goods with digitalized and undigitalized tasks following Zeira (1998), Aghion *et al.* (2018) and Acemoglu and Restrepo (2018). Perfect competition in factor markets and optimal uses of them lead to aggregate production function of final goods. The structure of model is depicted in the flow chart diagram in Figure 5.

[6]Some recent General AI technologies such as ChatGPT may even have an asymmetric influence on jobs where it replaces more skilled jobs (such as programmers, financial market brokers and analysts, etc.) rather than unskilled jobs (those reply more on physical tasks). Employing such asymmetric change in distribution of worker ability will make the result of our second feature even more significant.

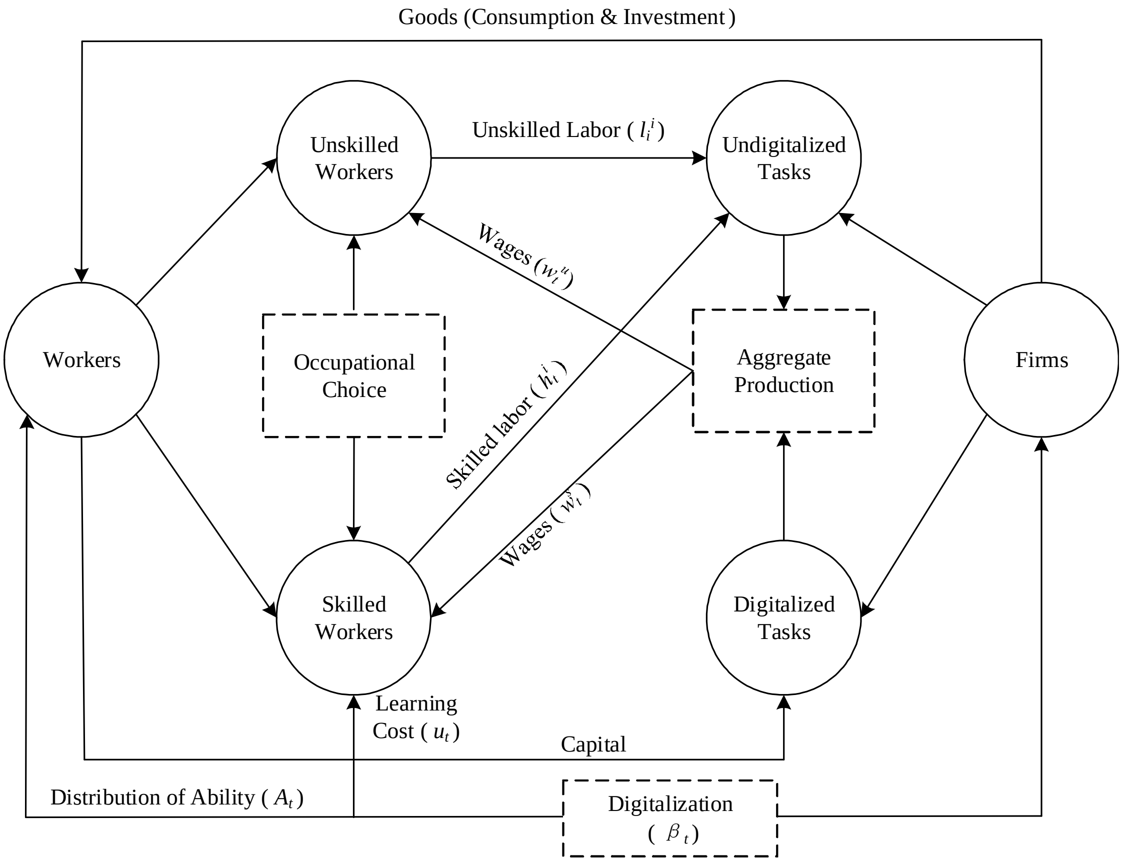

Figure 5: Model Structure

3.2 Digitalization and Firm Production

We assume that competitive firms produce final goods Y_t by combining different intermediate goods Y_{it} manufactured through different tasks via a constant elasticity of substitution (CES) production function:

$$Y_t = \widetilde{A} \left(\int_0^1 Y_{it}^{\rho} di \right)^{1/\rho} \tag{1}$$

where $\widetilde{A}$ is exogenous total factor productivity or common knowledge, $\frac{1}{1-\rho}$ captures elasticity of substitution across intermediate goods produced by different tasks. When $\rho < 0$, we have $\frac{1}{1-\rho} < 1$ thus different intermediate goods are complements rather than substitutes.

Each task has two options of production:

$$Y_{it} = \begin{cases} L_{it}, \text{ for un-digitalized task} \\ Z_t K_{it}, \text{ for digitalized task} \end{cases} \tag{2}$$

where L_{it} is composite labor supplied by both skilled and unskilled workers, K_{it} is capital used in production for digitalized task and Z_t represents a capital-augmenting technology. Examples of Z_t are technologies to create better machines, better digital devices and platforms, higher computer power, etc. The way Z_t enters the production function for digitalized task allows for the existence of balanced growth path (shown below), following Grossman *et al.* (2017) and Aghion *et al.* (2018).

Market clearing of factors market requires that:

$$\int_0^1 K_{it} di = K_t$$

$$\int_0^1 L_{it} di = L_t$$

$$Y_t = C_t + I_t = (1 - s) Y_t + sY_t$$

Define β_t as the rate of digitalization, it is shown that when factors are optimally utilized, the aggregation (1) can be written as

$$Y_t = \widetilde{A} \left[\beta_t \left(\frac{Z_t K_t}{\beta_t} \right)^{\rho} + (1 - \beta_t) \left(\frac{L_t}{1 - \beta_t} \right)^{\rho} \right]^{1/\rho} \tag{3}$$

Therefore, there are two forms of factor-augmenting technology in the economy, one is capital-augmenting technology Z_t, the other is digitalization. As noted by Aghion *et al.* (2018), this social production has the property that digitalization β_t is capital-depleting while labor-augmenting. This point is evident if we rewrite it in Cobb-Douglas type form:

$$Y_t = A_t \left[(Z_t B_t K_t)^{\rho} + (C_t L_t)^{\rho} \right]^{1/\rho} \tag{4}$$

with $B_t = (\beta_t)^{\frac{1-\rho}{\rho}}$ and $C_t = (1 - \beta_t)^{\frac{1-\rho}{\rho}}$. Because $0 < \beta_t < 1$ and $\rho < 0$, B_t actually decreases in β_t while C_t increases in β_t. This is why Z_t is introduced, i.e., it offers the opportunity to augment capital to offset the depleting effect of digitalization to deliver a BGP.

The shares of capital and labor income in total output are given by:

$$\alpha_{K_t} = \frac{MPK_t \cdot K_t}{Y_t} = \widetilde{A}^{\rho} \beta_t^{1-\rho} \left(\frac{Z_t K_t}{Y_t} \right)^{\rho} \tag{5}$$

$$\alpha_{L_t} = \frac{MPL_t \cdot L_t}{Y_t} = \widetilde{A}^{\rho} (1 - \beta_t)^{1-\rho} \left(\frac{L_t}{Y_t} \right)^{\rho} \tag{6}$$

respectively. Thus the ratio of the two share is

$$\frac{\alpha_{K_t}}{\alpha_{L_t}} = \left(\frac{\beta_t}{1 - \beta_t} \right)^{1-\rho} \left(\frac{Z_t K_t}{L_t} \right)^{\rho}. \tag{7}$$

PROPOSITION 1 *(Existence of balanced growth path): For* $0 < \beta_t < 1$, $n = 0$ *and* $\rho < 0$, *the economy achieves balanced growth path if either:*

a) $g_{Z_t} = 0$ *and* $\beta_t \longrightarrow 1$*, which is an asymptotic BGP (we name it BGP 1 in what follows) that delivers* $g_{k_t} = g_{y_t}$*;*

b) Or, $g_{Z_t} = 0$ *and* $\beta_t < 1$*, which is a special case of zero growth economy:* $g_{k_t} = g_{y_t} = 0$*;*

c) Or, $g_{Z_t} > 0$ *and* $g_{Z_t} = g_{k_t}(1-\beta_t)/\beta_t$*, which delivers is a BGP (we name it BGP 3 in what follows) with* $g_{k_t} = g_{y_t}$*.*

Proof of PROPOSITION 1: See Appendix A.1.

Discussion.

Case b) implies an economy without growth, thus is discarded.

Case a) and BGP 1 requires minimal assumption about capital-augmenting technology (i.e., a constant Z_t), e.g.,

$$Z_t = 1$$

while Case c) and BGP 3 amounts to imposing a endogenous process for Z_t,

$$Z_t = e^{t \cdot \psi \cdot g_{Z_t}}, \text{ with } g_{Z_t} = g_{k_t}(1-\beta_t)/\beta_t \tag{8}$$

such that it moves in the opposite direction of digitalization β_t.

Therefore, Case c) is a general case for Case a) (By setting $\psi = 0$, we move back to to Case a). In what follows, we will focus on Case c) and BGP 3 take it as our benchmark case.

On BGP, the growth rates are given by:

$$\begin{aligned} g_{Z_t} &= \left(\frac{\rho-1}{\rho}\right) g_{\beta_t}. \\ g_{k_t} &= g_{y_t} = \left[\frac{(\rho-1)\beta_t}{\rho(1-\beta_t)}\right] g_{\beta_t}. \end{aligned}$$

One example of steady state growth is $g_{\beta_t} = \frac{(1-\beta_t)}{\beta_t}\theta$, i.e., a constant fraction of undigitalized tasks is digitalized. Put differently, as more tasks are digitalized, the growth rate of digitalization slows down. Under this specification, the BGP is given by $g_{k_t} = g_{y_t} = \left(\frac{\rho-1}{-\rho}\right)\theta$.

3.3 SBTC and Labor Demand

We follow the literature and first analyze how digitalization affect labor demand and wages. This has been the central topic since 1990s. To analyze possible asymmetric effects of digitalization on different groups of workers, we introduce heterogeneity of labor force following Galor and Moav (2000) and divide labor supply into skilled and unskilled workers.

Labor demand. Assume that workers are free to choose to become skilled or unskilled workers at time t, thus the total labor force is a combination of them. However, the labor demand is skill-biased under digitalization, i.e., firms total employment is given by:

$$L_t = \gamma h_t + (1 - \delta \beta_t)\, l_t \tag{9}$$

where $\gamma > 1$ represents the higher preference for skilled labor and $0 < \delta < 1$ captures the discrimination of unskilled labor by firms when digitalization is present. The above setting of aggregate labor composition reflects the literature on skill-biased technical change.

Wages. The general wage paid to labor is given by the marginal product of labor:

$$w_t = MPL = \widetilde{A}^{\rho} \left[(1 - \beta_t) \frac{Y_t}{L_t} \right]^{1-\rho} = \widetilde{A}^{\rho} w\left(k_t, \beta_t\right) \tag{10}$$

with $y_t = \frac{Y_t}{L_t}$, $k_t = \frac{K_t}{L_t}$.

Given the composition labor in (9), the wage paid to a skilled worker is given by:

$$w_t^s = \gamma \widetilde{A}^{\rho} w\left(k_t, \beta_t\right) \tag{11}$$

and the wage paid to an unskilled worker is given by

$$w_t^u = (1 - \delta \beta_t)\, \widetilde{A}^{\rho} w\left(k_t, \beta_t\right) \tag{12}$$

3.4 Ability, Learning Cost and Labor Supply

It is well known that the above specification of production function usually implies monotonic positive relationship between digitalization and wage income gap. The shortcoming of this prediction however, is that it relies solely on labor demand effect of digitalization whilst the labor supply effect is completely ignored. To have a comprehensive analysis of the labor market outcomes, we extend the model with labor supply and human capital accumulation. After that, two novel features of labor supply are introduced to the model to capture how digitalization affects labor supply and the dynamics of wage gap: i) an erosion effect of digitalization on the distribution of worker abilities, and ii) a learning cost for workers who intend to become skilled ones.

Ability, learning cost and human capital accumulation. Both skilled and unskilled workers can accumulate human capital to increase h_t and l_t. They differ however, in endowed abilities and time left for working due to learning. Firstly, the unskilled workers accumulate human capital according to:

$$l_t^i = 1 - \left(1 - a_t^i\right) \beta_t \tag{13}$$

whilst skilled workers accumulate human capital according to:

$$h_t^i = (1-\tau)\left[a_t^i - \left(1-a_t^i\right)\beta_t\right] \tag{14}$$

where a_t^i is realized abilities from a uniform distribution between the region $a_t^i \in [A_t - 1, A_t]$. When a worker chooses to become a skilled one, he/she must spend a fraction of time, τ to learn high skills that complement with digitalization, so the working time reduces to $(1-\tau)$. Comparing (13) and (14), we see that both workers human capital are eroded by digitalization by $(1-a_t^i)\beta_t$. However, ability a_t^i can partially offset this erosion effect by extent of $a_t^i\beta_t$ - a augmenting effect. In addition, skilled workers may or may not have more human capital than unskilled workers, which depends on specific realized value of a_t^i. Thus, the SBTC setting in last subsection still plays an important role of creating a possible wage income gap.

Next, we make two novel extensions of human capital accumulation process:

i) The distribution of worker ability (A_t) evolves with progress in both common technology and digitalization:

$$A_t = \widetilde{A} - \widetilde{\phi}(\beta_t), \tag{15}$$

That is, improvements in general knowledge $\widetilde{A}$ equip workers higher abilities to learn, whilst digitalization make both workers harder to catch up with required abilities to complement. The latter is the creative destruction effect of digitalization. One might think that digitalization reduces ability of unskilled workers while benefits skilled workers. However, this turns out not be true for digitalization. For example, recent work by Brynjolfsson et. al. (2023) find evidence that the use of generative AI disseminates the tacit knowledge of more skilled workers, making them more substitutable and thus less important to firms. For another example, digitalization may also force high skilled workers to take low-skill jobs if their high-skill jobs are replaced. China now has near 100 million take-away riders, many of whom were white-collar workers with bachelor or higher degrees[7].

It is noted that $\widetilde{\phi}(\beta_t)$ is a general form. There are alternative ways to pin down a specific form for it. One way is to specify $\widetilde{\phi}(\beta_t)$ as a function of capital and labor shares

$$A_t = \widetilde{A} - \varphi\left(\alpha_{K_t}/\alpha_{L_t}\right), \ \varphi > 0 \tag{16}$$

That is, if digitalization raises $\alpha_{K_t}/\alpha_{L_t}$ (e.g., the case of an asymptotic BGP 1 in Proposition 1 a), or BGP 3 in Proposition 1 but the economy starts from a lower capital share

[7] An alternative treatment would be assuming that as digitalization deepens, they hurt high-skilled workers even more than low-skilled workers, thus imposing asymmetric ability erosion effects to two groups of workers. Whilst this may be possible for latest digital technologies such as the adoption of a generative AI, we discard this alternative treatment for two reasons. The first reason is that it is hard to judge the degree of asymmetry. Secondly, workers in most of our sample period did not see the latest development of generative AI.

state), it reduces the importance of labor as a whole as a result of substitution effect or as it causes workers loose part of their abilities during unemployment. Since factor shares are ultimately functions of β_t and $Z_t K_t$, adoption of this form requires tracking dynamic changes in these variables before evaluating how they affect worker ability distribution A_t.

A simpler option is to follow a linear rule

$$A_t = \widetilde{A} - \phi\beta_t, \ \phi > 0 \tag{17}$$

Similar simple rule was used in Acemoglu and Restrepo (2018) where they specify an evolution rule for productivity when there is automation technology $I(t)$: $n_t = N_t - I_t$. That is, creation of new tasks, N_t (which is analogous to our general technology $\widetilde{A}$) raises productivity n_t while new automation technology I_t (which is analogous to our digitalization β_t) reduces it. We will consider both forms of (16) and (17) in numerical simulations, but only consider (17) when deriving analytical results. The parameter ϕ is a key parameter that governs the marginal impact of digitalization on worker ability distribution.

ii) Workers pay learning costs if they choose to become skilled ones:

$$\mu\left(w_t^s h_t^i, \beta_t\right) = w_t^s h_t^i \left(\mu_0 - \eta\beta_t\right). \tag{18}$$

This learning cost stems from the notion that workers have to pay learning costs before they become skilled ones. Examples of learning costs can be education costs, tuition fees and training costs for studying digital technology. Effectively, the existence of learning cost imposes transition costs (see, e.g., Cortes and Gallipoli 2018) on labor mobility and makes two groups of labor imperfect substitutes.

The learning cost is a function of the wage payments of high skilled workers, $w_t^s h_t^i$, and digitalization, β_t. The latter captures an important role of digitalization, i.e., workers get easier access to knowledge of new tasks, training programs and skill tutorials via digital platforms, and the costs of which are much lower (sometimes even free) than traditional training programs they pay on undigitalized platforms. In addition, the developments of AI and online tutorial resources could potentially make workers learn faster than before. This kind of positive effect of digitalization on worker ability is analogous to positive externality/nonrivalry property of data economics. But they are mainly adopted in firm production behavior. We apply this positive learning effect in labor market and analyze its direct effect on labor supply. Notably, the parameter η governs the size of learning effect.

3.5 Labor Income and Occupational Choice

Wage incomes of two groups. Income of skilled workers net of learning cost:

$$\begin{aligned} I_t^{i,s\_net} &= I_t^{i,s} - \mu\left(w_t^s h_t^i, \beta_t\right) \\ &= w_t^s h_t^i - \mu\left(w_t^s h_t^i, \beta_t\right) \end{aligned}$$

and

$$\begin{aligned} w_t^s h_t^i &= \gamma\left(1-\tau\right) \widetilde{A}^{\rho} w\left(k_t, \beta_t\right) \left[a_t^i - \left(1-a_t^i\right)\beta_t\right] \\ &= \widetilde{A}^{\rho} w\left(k_t, \beta_t\right) \left[a_t^i - \left(1-a_t^i\right)\beta_t\right] \end{aligned}$$

That is, for simplicity we have assumed that $\gamma\left(1-\tau\right)=1$. So, we have:

$$I_t^{i,s\_net} = \widetilde{A}^{\rho} w\left(k_t, \beta_t\right) \left[a_t^i - \left(1-a_t^i\right)\beta_t\right] - \widetilde{A}^{\rho} w\left(k_t, \beta_t\right)\left(\mu_0 - \eta\beta_t\right).$$

Threshold (division of skilled v.s. unskilled workers):

Workers will choose to be skilled workers (get education/learning/training and pay cost) if and only if:

$$I_t^{i,s\_net} \geq I_t^{i,u} \tag{19}$$

The threshold then is derived when the incomes of two options take equal sign:

$$\begin{aligned} & \widetilde{A}^{\rho} w\left(k_t, \beta_t\right) \left[a_t^i - \left(1-a_t^i\right)\beta_t\right] - \widetilde{A}^{\rho} w\left(k_t, \beta_t\right)\left(\mu_0 - \eta\beta_t\right) \\ = \; & \left(1-\delta\beta_t\right) \widetilde{A}^{\rho} w\left(k_t, \beta_t\right) \cdot \left[1 - \left(1-a_t^i\right)\beta_t\right] \end{aligned}$$

$$\begin{aligned} & \left[a_t^i - \left(1-a_t^i\right)\beta_t\right] - \left(\mu_0 - \eta\beta_t\right) \\ = \; & \left(1-\delta\beta_t\right)\left[1 - \left(1-a_t^i\right)\beta_t\right] \end{aligned}$$

which solves for a_t^*:

$$a_t^* = \frac{\left(1-\delta\beta_t+\delta\beta_t^2\right) + \left(\mu_0 - \eta\beta_t\right)}{\left(1+\delta\beta_t^2\right)} \tag{20}$$

PROPOSITION 2 *(Occupational choice): Given meaningful parameter values, workers occupational choice is determined by the threshold in (20). Workers whose abilities above this threshold choose to be skilled workers, the remaining workers choose to be unskilled workers. And the threshold has the following properties:*

a) The threshold, a_t^* is independent with the distribution of worker ability A_t.

b) The threshold, a_t^* is decreasing in learning cost, $\mu\left(w_t^s h_t^i, \beta_t\right)$ for $\eta > 0$.

c) The threshold, a_t^* is decreasing in digitalization, β_t for $\eta > 0$.

Proof of PROPOSITION 2: See Appendix A.2.

Discussion. The second term of Proposition 2 implies that the introduction of linear and negatively associated learning cost to digitalization induce a linear relationship between the threshold and digitalization. While this is a straightforward result, one may argue for alternative assumptions of the function forms of learning cost. Arguably, one possible specification is a quadratic cost function for learning cost (or for worker ability function). We do not consider this alternative quadratic form here in the benchmark for two reasons. First, a quadratic functional form is ad hoc and hardly identifiable by evidence, and second, such quadratic form self-ensures nonlinear relationship between wage income gap and digitalization and thus is too restrictive (see Appendix A.2 for a demonstration).

3.6 Labor Income Inequality

We explore wage income inequalities that are *observable* in data: $i^S(\cdot)$ and $i^U(\cdot)$. The variations in learning costs are not directly observable in real world. Then, we compute three measures of these observable income inequalities: average wage income inequality *between* two groups of workers, *within-group* inequalities of skilled workers and *within-group* inequality of unskilled workers, and examine how they are affected by changes in digitalization β_t in the presence of labor demand effect of SBTC and two labor supply effects discussed above.

3.6.1 Between-group inequality

First of all, the *average* income of skilled workers is given by:

$$\begin{aligned}\widetilde{I}_t^S &= \frac{i^S(A_t) + i^S(a_t^*)}{2} \\ &= W_t \frac{\overbrace{a_t^*}^{\text{Labor supply effect}} + A_t - \overbrace{(2 - A_t - a_t^*)\beta_t}^{\text{Labor demand effect } a_t^*\beta_t}}{2}\end{aligned}$$

In a similar way, average income of unskilled workers is given by:

$$\begin{aligned}\widetilde{I}_t^U &= \frac{i^U(A-1) + i^U(a_t^*)}{2} \\ &= W_t(1-\delta\beta_t)\frac{2 - (3 - A_t - a_t^*)\beta_t}{2}\end{aligned}$$

Therefore, *between-group* wage inequality can be obtained:

$$\sigma_t^{\frac{S}{U}} = \frac{\widetilde{I}_t^S}{\widetilde{I}_t^U} = \frac{a_t^* + A_t - (2 - A_t - a_t^*)\beta_t}{(1-\delta\beta_t)\left[2 - (3 - A_t - a_t^*)\beta_t\right]}. \tag{21}$$

Given the above formulas of various measures of wage income inequalities, we can see that those inequality measures are not only a direct function of digitalization β_t, but also functions of the worker ability upper bound A_t and the threshold of worker ability a_t^*. Since A_t and a_t^* are also functions of digitalization (see equation 15 and 20), those wage income inequalities are ultimately a function of digitalization β_t. Based on these observations, we propose proposition 3.

PROPOSITION 3 *(Digitalization and between-group wage inequality): In this economy,*

a) Increases in digitalization, β_t always monotonically widen between-group wage inequality with fixed distribution of worker ability and frictionless labor mobility; (labor demand effect)

b) Ceteris paribus, increases in digitalization, β_t always monotonically shrink between-group wage inequality when worker ability distribution is shifted left by digitalization; (labor supply effect I)

c) Ceteris paribus, increases in digitalization, β_t always monotonically shrink between-group wage inequality when learning cost is negatively related to digitalization; (labor supply effect II)

d) The relationship between digitalization, β_t and between-group wage inequality is uncertain when the three effects coexist. An inverted U-shape emerges when different effects countervail each other in the short-run, depending on parameter values of $\{\delta, \phi, \eta\}$.

PROOF of PROPOSITION 3: See Appendix A.3.

Here presents some key elements of proof. Consider first the case that there is endogenous adjustment of worker ability ($\phi > 0$) but without learning cost ($\eta = 0$). We can examine the impact of digitalization β_t on between-group wage income inequality $\sigma_t^{\frac{S}{U}}$ by taking first order derivative of the latter to the former (see Appendix A.3).

$$\frac{\partial \sigma_t^{\frac{S}{U}}}{\partial A_t} = \frac{\overbrace{\left\{2 - \beta_t - \beta_t^2\right\}}^{\text{positive}}}{\left(1 - \delta\beta_t\right)\left\{\left[2 - \left(3 - A\left(\beta_t\right) - a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} > 0.$$

Since

$$A'\left(\beta_t\right) = -\phi < 0$$

It is clear that introducing creative destruction effect (endogenizing A_t) helps to reduce wage income inequality.

Next, consider the case that there is positive learning cost ($\eta > 0$) but no endogenous adjustment of worker ability ($\phi = 0$). In this case, only learning cost channel is in play. However, since the learning cost $\mu\left(w_t, \beta_t\right)$ only appears in the threshold a_t^*, we can easily check the relationship between $\sigma_t^{\frac{S}{U}}$ and learning cost $\mu\left(w_t, \beta_t\right)$ by examining the first derivative of $\sigma_t^{\frac{S}{U}}$ to a_t^*. It is easily verified that it is identical to the first derivative of $\sigma_t^{\frac{S}{U}}$

to A_t above. Since

$$a^{*\prime}(\beta_t) = \frac{-\left(1-\delta\beta_t^2\right)(\delta+\eta) - 2\delta\beta_t\mu_0}{\left(1+\delta\beta_t^2\right)^2} = a^{*\prime}(\beta_t;\delta,\eta,\mu_0) < 0$$

It is clear that introducing learning cost effect also helps to reduce wage income inequality.

Finally, to examine the overall effect of changes in β_t on $\sigma_t^{\frac{S}{U}}$, we have to derive $\frac{\partial\sigma_t^{\frac{S}{U}}}{\partial\beta_t}$ which generally will collect combined effects of labor demand effect of SBTC ($\delta > 0$), the creative destruction effect and learning cost effects on labor supply (for $\delta > 0, \eta > 0$). The Appendix A.3 shows that the derivation of $\frac{\partial\sigma_t^{\frac{S}{U}}}{\partial\beta_t}$ is rather tedious, and we will examine this numerically in the Simulation section.

$$\frac{\partial\sigma_t^{\frac{S}{U}}}{\partial\beta_t} = \frac{Num\left(\beta_t;\delta,\eta;\widetilde{A},\mu_0\right)}{\left(1-\delta\beta_t\right)^2\left[2-\left(3-A(\beta_t)-a^*(\beta_t)\right)\beta_t\right]^2} = f\left(\beta_t;\delta,\phi,\eta;\widetilde{A},\mu_0\right)$$

However, the principles are very clear: the ultimate relationship between digitalization and wage income inequality depends on the relative strengths of the labor demand effect of SBTC (enlarging wage gap) and the two labor supply effects (shrinking wage gap) when digitalization is ongoing. Those effects are governed by three key parameters $\{\delta,\phi,\eta\}$. In particular, there exists a combination $\{\delta^*,\phi^*,\eta^*\}$ that making $\frac{\partial\sigma_t^{\frac{S}{U}}}{\partial\beta_t} = 0$ which delivers an inverted U-shape with $\beta_t^*\{\delta^*,\phi^*,\eta^*\}$ as the inflection point.

3.6.2 Within-group inequality

Differently, *within-group* wage inequality is defined as the ratio between maximum wage income to minimum wage income in a particular group. Given this definition, the *within-group* of skilled workers is computed as:

$$\sigma_t^S = \frac{i^S(A_t)}{i^S(a_t^*)} = \frac{A_t-(1-A_t)\beta_t}{a_t^*-(1-a_t^*)\beta_t}. \tag{22}$$

Similarly, *within-group* wage inequality of unskilled workers is computed as:

$$\sigma_t^U = \frac{i^U(a_t^*)}{i^U(A_t-1)} = \frac{1-(1-a_t^*)\beta_t}{1-(2-A_t)\beta_t}. \tag{23}$$

PROPOSITION 4 *(Digitalization and within-group wage inequality): In this economy,*

a) Under the creative destruction channel, digitalization reduces worker ability, and the same time reduces within group inequality of skilled workers while raises within group inequality of unskilled workers.

b) Under the learning cost channel, digitalization reduces job threshold, and the same time raises within group inequality of skilled workers while reduces within group inequality of unskilled workers.

c) When all channels are active, the impact of digitalization on within group wage income inequality is uncertain and depends on $\{\beta_t; \delta, \phi, \eta\}$.

PROOF of PROPOSITION 4: See Appendix A.4.

Intuitions: For creative destruction channel, digitalization shifts work ability distribution to the left while the job choice threshold is unaffected (see Figure 6 below). This implies that the lowest wage earner of the skilled workers (and the highest wage earner of the unskilled workers) earn the same wage income as before. Therefore, the within group inequality of skilled workers must fall and the within group inequality of unskilled workers must increase. In contrast, the learning cost channel reduces costs of becoming skilled workers, lowers down job threshold, and thus increases the labor supply of skilled workers. Thus, the lowest wage earner of the skilled group could be the unskilled workers before, they enter the skilled group now but they drag down the lowest wage income of the skilled group. As a result, the within group inequality of skilled workers must rise and the within group inequality of unskilled workers must drop.

3.6.3 Understanding the effect of worker ability distribution

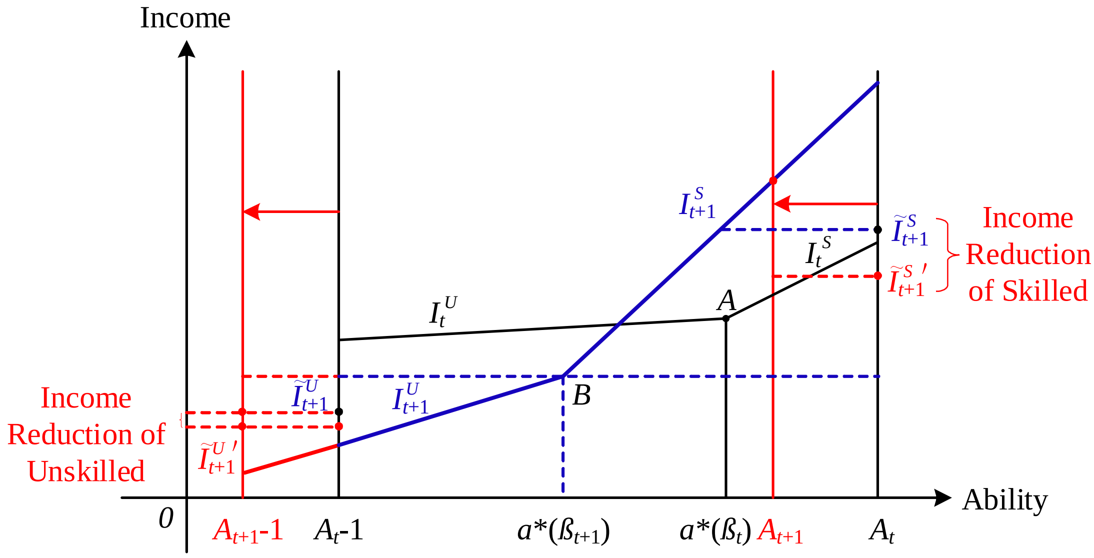

Figure 6: The effect of a fall in worker ability distribution

The implication of this workers overall ability ceiling evolution rule is that, other things equal, progress in digital technology reduces workers ability as a whole. Because this is a labor supply effect, it reduces wages of both groups. This put downward pressure on wage inequality as the reduction in wages of skilled workers is much more than those of unskilled workers. This can be depicted in Figure 6 where the black vertical lines represent the distribution of worker ability at time t and the two red vertical lines represent the distribution of worker ability at time t after digitalization evolves. We can see that, despite the fall of threshold thus more workers could have become skilled workers, the reduced level of worker ability distribution offsets the enlarging wage income inequality.

4 Empirical Evidence

4.1 Data

We constructed a panel dataset of China's digitalization with both provincial and industrial dimensions for the period 2007-2019 based on the "2021 Categorization of Core Industries of Digital Economy (CCIDE 2021). This Categorization has four major categories: digital products manufacturing, digital product services, digital technology application and data-driven industries. The first two categories focus on the new technologies in production of digital goods and services while the last two focus on digitalization of industries by adopting digital goods and services. As a result, it provides us a much wider measure of "digital technology'.

The data used for constructing digital economy measures comes from three main sources: Firstly, the input-output tables of each province for more than 100 sectors in 2007, 2012 and 2017. The input-output tables for more than 100 sectors have relatively detailed industry divisions and is the basis of statistical accounting for the digital economy, covering 21 provinces. Secondly, the 2008 and 2013 China Economic Census Yearbooks. Combining them with the input-output tables of each province, the value added of the digital economy in each province can be calculated. Thirdly, the statistics at the provincial level, including indicators such as skill premium, GDP per capita, fixed capital stock per capita and urbanization rate, are mainly from provincial and municipal statistical yearbooks and the China Statistical Yearbook.

Table 1: Description of Variables

Variables	Formula of Construction
premium1	Wage income ratio between SRTS and manufacturing
premium2	Wage income ratio between SRTS and agriculture
digit	Proportion of overall value added of digital economy
digit^2	The square of digit
perk	Capital stock/Resident population
gdp	Gross production/Resident population
urban	Urban employed population/Total population
cost	Consumer spending/Disposable income
industryad	Tertiary sector output/Secondary sector output
industryra	Tertiary sector output/GDP
social	Social security expenditure/General budget expenditure
edu	Education expenditure/General budget expenditure
fdi	Total investment in foreign-invested enterprises/GDP
imandex	Total imports and exports/GDP
state	Employees in state-owned enterprises/total employment

The constructed data variables are described in Table 1. The dependent variable is measured by the ratio of the average wage in Scientific Research and Technical Services (SRTS) to the average wage in manufacturing (premium1). As a robustness check, an alternative measure is the ratio of the average wage in Scientific Research and Technical Services (SRTS) to the average wage in agriculture (premium2).

The key explanatory variable is the proportion of value added of digital economy to local GDP in each province in China. According to the methodology of the US Bureau of Economic Analysis (BEA), the value added of digital economy at the provincial and municipal levels is statistically accounted for using input-output tables, statistical yearbooks and economic census data in 21 provinces, based on the Catalogue of Core Industries of the Digital Economy (2021) published by the National Bureau of Statistics. The accounting is divided into four core categories: digital product manufacturing, digital product services, digital technology applications and data-driven industries. In addition, 11 variables were selected as control variables from five perspectives, reflecting five aspects: level of economic development, structural changes in the economy, government behavior, international trade and the marketization process.

4.2 Descriptive Statistics

The descriptive statistics of the variables used in measurement and estimation are described in Table 2. Table 3 also shows the sizes of digital economy and their shares to GDP in selected provinces and municipalities in 2017.

Table 2 Descriptive Statistics of Variables

Variables	Obs	Mean	Std. Dev	Min	Max
premium1	434	2.337	0.793	1.267	6.881
premium2	434	1.482	0.276	0.949	3.12
digit	231	0.07	0.034	0.026	0.191
$digit^2$	231	0.006	0.006	0.001	0.037
perk	330	17.37	9.433	2.952	56.403
gdp	434	10.566	0.566	8.959	12.009
urban	434	0.554	0.142	0.215	0.896
cost	434	0.736	0.055	0.56	0.905
industryad	434	1.251	0.688	0.527	5.244
industryra	434	0.476	0.093	0.298	0.837
social	434	0.13	0.035	0.055	0.276
edu	434	0.163	0.026	0.099	0.222
fdi	434	0.505	1.837	0.047	37.212
imandex	434	0.289	0.325	0.008	1.587
state	434	0.331	0.135	0.083	0.842

Table 3: Sizes of Digital Economy in Selected Provinces and Municipalities in 2017

	Digital Product Manufacturing	Digital Product Services	Digital Technology Application	Data-driven Industries	Share to GDP
Beijing	428.6585	70.95416	3666.653	200.9252	14.61
Tianjin	487.7475	60.5429	520.6288	55.8851	9.03
Zhejiang	1770.319	165.9525	2716.879	182.2212	9.22
Jiangsu	5324.442	211.7899	3012.713	316.8356	10.32
Fujian	1295.28	62.79458	820.7065	76.24486	6.66
Guangdong	8388.932	253.0411	3913.783	205.0231	13.92
Liaoning	478.6626	52.17442	577.1305	48.5994	5.33
Jilin	125.794	91.38116	361.2722	54.60505	5.79
Heilongjiang	105.4805	23.63477	215.2625	19.60313	2.95
Sichuan	1358.679	90.07507	1687.829	59.7422	8.43
Guizhou	169.7666	27.29571	380.6914	44.4132	4.57
Yunnan	171.6392	50.94098	504.3869	67.17601	4.29

Before performing a formal econometric analysis of variables, it is useful to have a intuitive perception of the wage gap and digitalization relationship. Figure 7 shows a scatter plot of the two variables at their average values over the sample time periods 2017-2019. It looks more like a inverted U-shape rather than a monotonic relationship. This is at odds with the current theoretical predictions in the literature. The following will test this relationship using panel data regression formally.

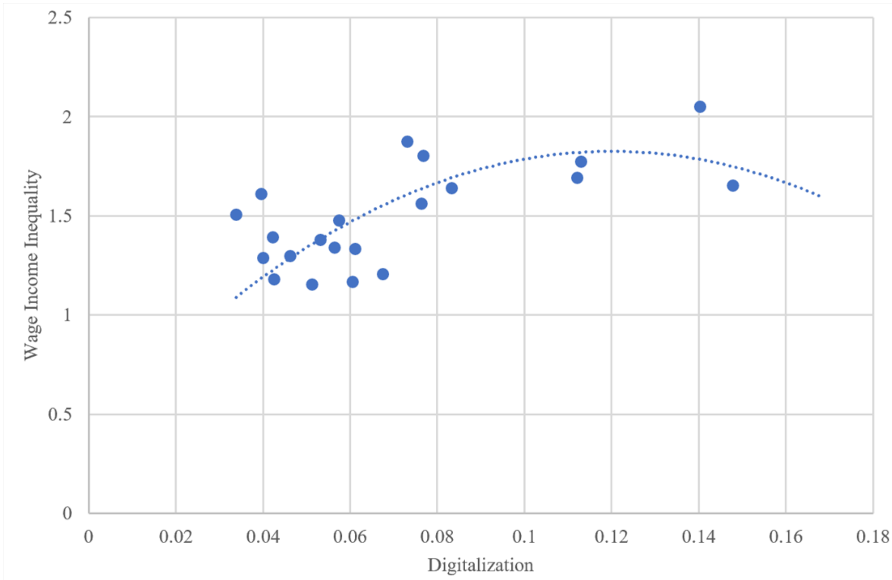

Figure 7: China's digitalization and wage income inequality - A simple scatter plot.

4.3 Empirical Model

Combined with the previous theoretical analysis, we first develop an econometric model between the digital economy and the skill premium as:

$$premium_{i,t} = \alpha_0 + \alpha_1 digit_{i,t} + \alpha_2 digit_{i,t}^2 + \alpha_3 X_{i,t} + \alpha_t + \alpha_i + \varepsilon_{i,t} \tag{24}$$

where $premium_{i,t}$ denotes the skill premium (measured either by premium1 or premium2), $digit_{i,t}$ denotes the proportion of digital economy to local GDP (a measure of digitalization), and $X_{i,t}$ denotes a set of control variables that affect the skill premium. The subscripts i and t denote province and year respectively; α_i denotes unobservable province fixed effects, with individual fixed effects added to control for province characteristics; α_t denotes time fixed effects, with time fixed effects added to control for year-specific event effects; and $\varepsilon_{i,t}$ is the error term. Meanwhile, we analyze the relationship between the development of the digital economy and the skills premium at the district level, so the standard errors are clustered to the district level. According to model (1), α_1 and α_2 measure the overall impact of the development of the digital economy on the skills premium.

To examine the transmission mechanism of the main relationship from digitalization to wage premium, we also test mediating effect by setting up two additional regressions:

$$growth_{i,t} = \beta_0 + \beta_1 digit_{i,t} + \beta_2 digit_{i,t}^2 + \beta_3 X_{i,t} + \alpha_t + \alpha_i + \varepsilon_{i,t} \tag{(2)}$$

$$premium_{i,t} = \gamma_0 + \gamma_1 digit_{i,t} + \gamma_2 digit_{i,t}^2 + \gamma_3 growth_{i,t} + \gamma_4 X_{i,t} + \alpha_t + \alpha_i + \varepsilon_{i,t} \tag{(3)}$$

where growth$_{i,t}$ is our constructed mediator which denotes the speed of digital technology progress. It is obtained as the principal components of four selected indicators: total telecommunication services per capita, internet broadband access subscribers, number of mobile phone subscribers and IT service revenue. The growth rate of this index is then taken as a proxy to represent the speed of digital technology progress and thus a gauge of learning effects or effects of worker ability change. The mediating effect can be tested statistically using the three-step approach developed by Baron and Kenny (1986).

4.4 Empirical results

4.4.1 Main results

The benchmark results for the impact of the digital economy on the skills premium are given in Column (1) and (2) in Table 4. The results in column (2) show that all other factors being equal, the coefficient of the primary term of the digital economy on the skill premium is 10.86 and the coefficient of the squared term of the digital economy on the skill premium is -67.77, both passing the significance test. In other words, the digital economy has an "inverted U" shaped effect on the skill premium. The inflection point is thus around 8%.

Table 4: Panel regression results

Variable	(1) premium1	(2) premium1	(3) premium1	(4) premium1	(5) premium2	(6) premium2
			(IV)	(IV)	(IV)	(IV)
digitprop	-1.65	10.86***	-0.29	12.69***	7.45*	34.16***
	(-0.93)	(2.75)	(-0.21)	(4.57)	(2.56)	(5.93)
digitprop^2		-67.77***		-69.61		-143.2***
		(-3.14)		(-5.12)		(-5.13)
Control variables	YES	YES	YES	YES	YES	YES
Year effect	YES	YES	YES	YES	YES	YES
Provincial effects	YES	YES	YES	YES	YES	YES
LM statistic					42.864	44.73
					(0.00)	(0.00)
Wald F statistic					656.56	322.18
					(16.38)	(7.03)
N	231	231	210	210	210	210
adj-R^2	0.51	0.56	0.51	0.56	0.42	0.48

t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Next, we conduct two robustness checks. The first robustness check involves internal instrumental variables. Although the introduction of year and province as control variables in the previous model can solve the endogeneity problem to a certain extent, it cannot completely avoid the endogeneity problem caused by omitted variables and so on. Based on this, the one-period lag of the digital economy is used as the instrumental variable. The one-period lagged variable, as a traditional instrumental variable, is able to meet the requirements of relevance and exclusivity. Columns (3) and (4) of Table 4 give the regression results for the instrumental variables. Firstly, the results of the under-identification test (Kleibergen-Paap rk LM statistic test) as well as the weak identification test (Cragg-Donald Wald F statistic test) for the instrumental variables prove the validity of the instrumental variables. Secondly, the empirical results suggest that digital economy still shows an inverted U-shaped relationship with the skill premium and is more significant.

The second robustness check is to replace the value added of the digital economy with the share of the digital economy in GDP (prop). The results in columns (5) and (6) of Table 4 show that the share of the digital economy in GDP still shows a significant inverted U-shaped relationship with the skill premium.

4.4.2 Mediating Effect

The results of the test for the mediation effect are shown in Table 5. Column (2) shows that the coefficient of impact of the digital economy primary term (digit) on the speed of digital skill progress (growth) is 3.918, and the coefficient of impact of the digital economy squared term ($digit^2$) on the speed of digital skill progress (growth) is -0.296, and both pass the significance test at the 1% level, in other words, the relationship between the digital economy and the rate of digital skill progress shows an inverted U-shaped curve. In column (3), the coefficient of the impact of the speed of digital technological progress on the skill premium is 0.262 and passes the significance test at the 5% level, indicating that the faster the speed of digital technological progress, the stronger the advantage of high-skilled labour, and then the greater the skill premium. Therefore, the mediating effect of the speed of digital technology progress is significant.

Table 5: Mediating effect

Variable	(1) premium1	(2) growth	(3) premium1
digit	1.363***	3.918***	-1.074
	(5.64)	(5.25)	(-2.02)
$digit^2$	-0.088***	-0.296***	0.0863
	(-6.05)	(-5.91)	(2.32)
growth			0.262**
			(3.17)
Control variables	YES	YES	YES
Instrumental variable	YES	YES	YES
Year effect	YES	YES	YES
Provincial effects	YES	YES	YES
N	210	54	54
adj-R^2	0.6135	0.7223	0.7238
Intermediary effects		Significant	

t statistics in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Simulations

5.1 Calibration

Before simulating the model specified in last section, we discuss about calibration of parameters. The erosion of labor parameter δ takes value of 0.25. This is most commonly used value for example in the evolution of capital literature in macroeconomics. The natural rate of interest rate, or the long-run real interest rate is 2 percent which matches the average real interest rate in China. It is also close to values usually adopted for western countries. The parameter η which governs the size of learning cost is calibrated to 1 as the benchmark and it is set to 0 to mute off. It is hard to pin down this value from literature for the Chinese economy. Therefore, I follow the estimate of a similar parameter in western countries. We vary this parameter from 0 to 5 to check robustness. The parameter ϕ is the coefficient on work ability which captures the impact of digitalization on work ability upper bound. It can be set to 0 to mute off this channel or set to 1 to switch on. Finally, ψ is set to 1 for positive and capital-augmenting technology growth rate and to 0 for constant capital-augment technology. The calibration is summarized in Table 3.

Table 7: Parameter values

Parameters	Benchmark Value	Definition	Region
ρ	-1	Complementarity between tasks	$(-\infty, 0)$
δ_k	0.1	Depreciation of capital	$[0, 0.25]$
ψ	1	Coefficient in Z_t	1 *or* 0
δ	0.5	Depreciation of human capital	$[0, 0.9]$
s	0.2	Saving rate	$[0.1, 0.5]$
$\widetilde{A}$	1	Steady state common technology	$(0, +\infty)$
μ_0	0	Fixed learning cost	$[-1, 1]$
η	0.66	Coefficient of learning cost	$[0, +\infty]$
ϕ	0.66	Coefficient of work ability	$[0, +\infty]$

5.2 Digitalization, labor demand effect and inequality

We first simulate the model given exogenous progress of digitalization when only SBTC channel is active. That is, we use fixed distribution of worker ability ($\phi = 0$) and zero learning cost ($\eta = 0$). This is to set out a benchmark which the later simulations can compare to. The result is shown in Figure 8.

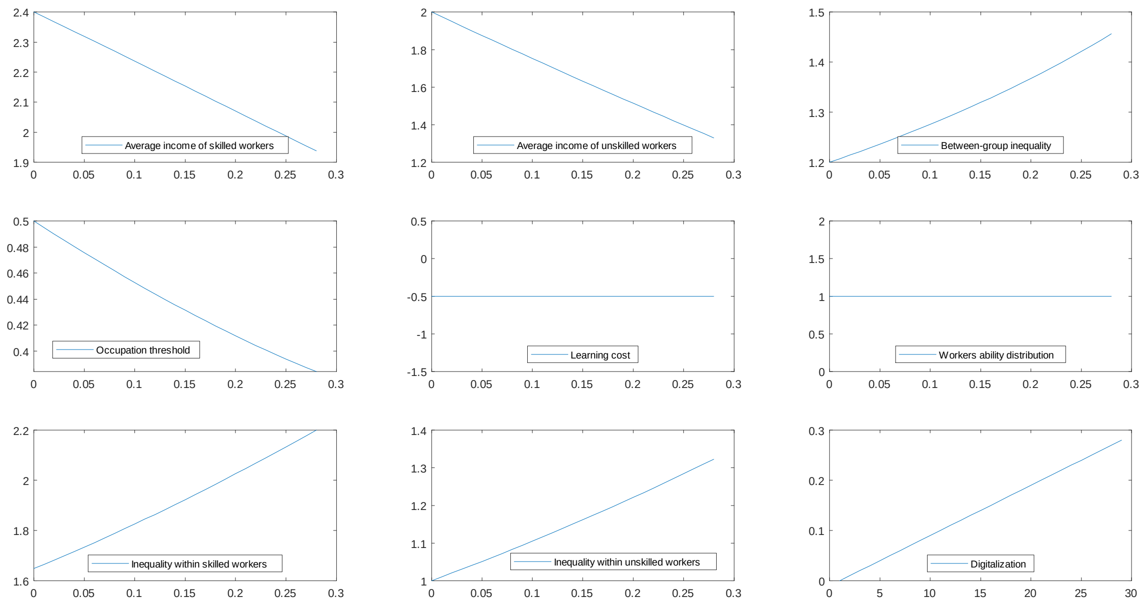

Figure 8: Model simulation with only labor demand effect ($\delta = 0.5$, $\phi = \eta = 0$).

It is seen that the model generate monotonic and increasing relationship between digitalization and wage income inequality (both between-group and within-group). This is consistent with the findings in Galor and Moav (2000), our Proposition 3 a) and is also in line with the majority of literature on the impacts of ICT technology on wage income gap. In addition, the within-group wage income gap show positive relationships with digitalization for both skilled and unskilled workers. However, these results are obtained under the assumption that digitalization only affects labor demand in the labor market.

5.3 Digitalization, labor ability distribution and inequality

We then allow for labor supply effects of digitalization. The first novel mechanism we introduce to labor supply side is that digitalization exerts a creative destruction effect on ability of both workers ($\phi > 0$). To get a clean comparison, we still assume zero learning cost ($\eta = 0$). The results are shown in Figure 9 ($\phi = 0.7$) and Figure 10 ($\phi = 1.4$).

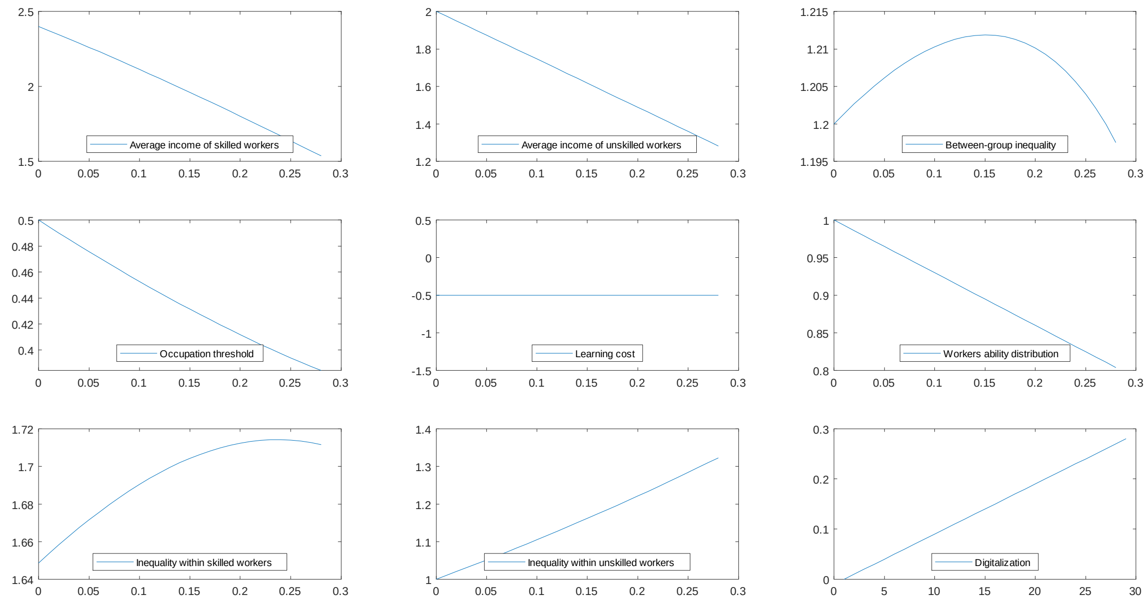

Figure 9: Model simulation with labor demand effect and endogenous work ability distribution ($\delta = 0.5$, $\phi = 0.7$, $\eta = 0$).

Compared with Figure 8, we now allow for two countervailing forces to wage income gap. One is the SBTC force that raises (reduces) labor demand of skilled (unskilled) workers. This traditional force enlarges wage income gap along digitalization. The other force is the creative destruction force that shrinks the ability distribution of all workers which reduces wage income gap. Figure 9 just capture the state that those two countervailing forces offset each other and creates an inverted U-shape (see top-right panel of Figure 9). This is consistent with our theoretical prediction in Proposition 3 d). In addition, the within-group wage income gaps still show positive relationships with digitalization for both skilled and unskilled workers. However, the within-group wage gap of skilled workers starts to become concave (see bottom left panel of Figure 9). This is a result of combining labor demand and labor supply effects (consistent with Proposition 4 b)), although the latter effect is not strong enough to dominate the former.

Of course, the two countervailing forces may run into imbalance. One example is that the creative destruction effect dominates. This case is plotted in Figure 10 where the elasticity ϕ is doubled. As a result, the between-group inequality displays a negative monotonic shape, consistent with Proposition b). In addition, the within-group inequality of unskilled workers are the same as in Figure 9, but the same of skilled workers changes to be negatively slopped (see bottom left panel of Figure 10). This is again consistent with

Proposition 4 c).

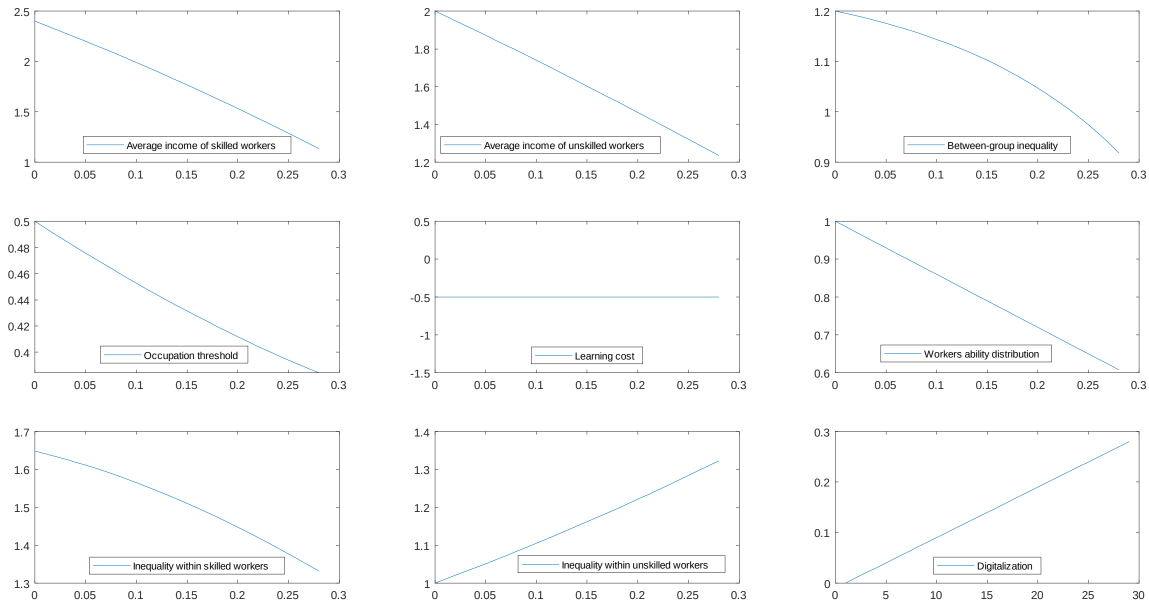

Figure 10: Model simulation with labor demand effect and endogenous work ability distribution ($\delta = 0.5$, $\phi = 1.4$, $\eta = 0$).

5.4 Digitalization, learning cost and inequality

We now fix worker ability distribution channel ($\phi = 0$) and switch on learning cost channel. As implied by the theoretical results, introducing learning costs for workers who intend to become skilled ones puts downward pressure on wage income inequality. The simulation results are shown in Figure 11 ($\eta = 1.35$) and Figure 11 ($\eta = 2.7$).

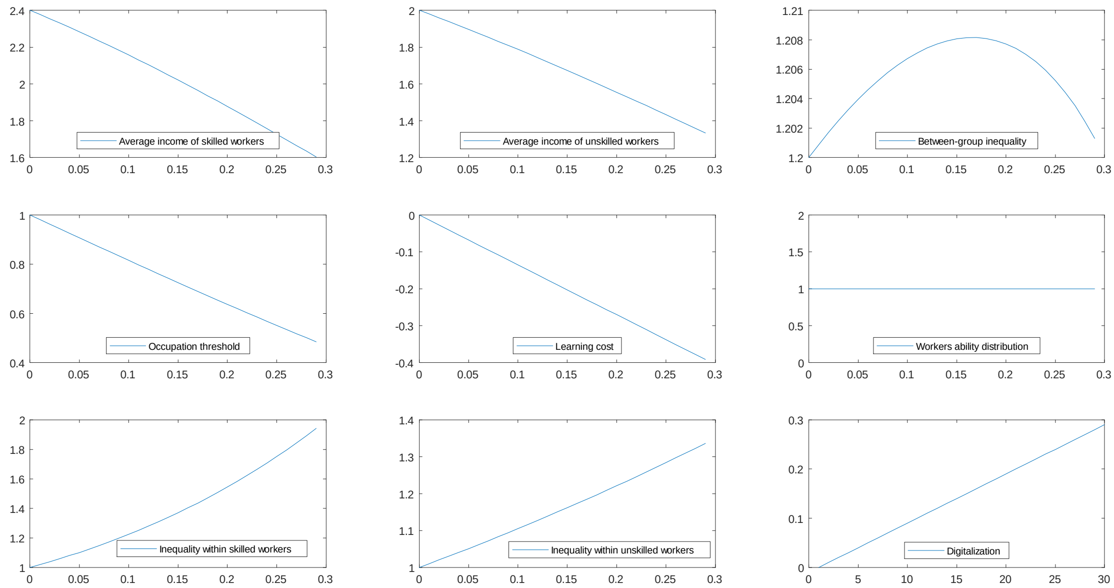

Figure 11: Model simulation with labor demand effect and endogenous learning cost ($\delta = 0.5$, $\phi = 0$, $\eta = 1.35$).

Figure 11 shows the particular case that labor supply effect of having learning cost just offsets the labor demand effect of SBTC. This result is similar with Figure 9 where creative destruction effect was introduced. The difference is that the within-group wage income gap of the skilled workers still show positive monotonic relationship. This result is consistent with Proposition 4 b).

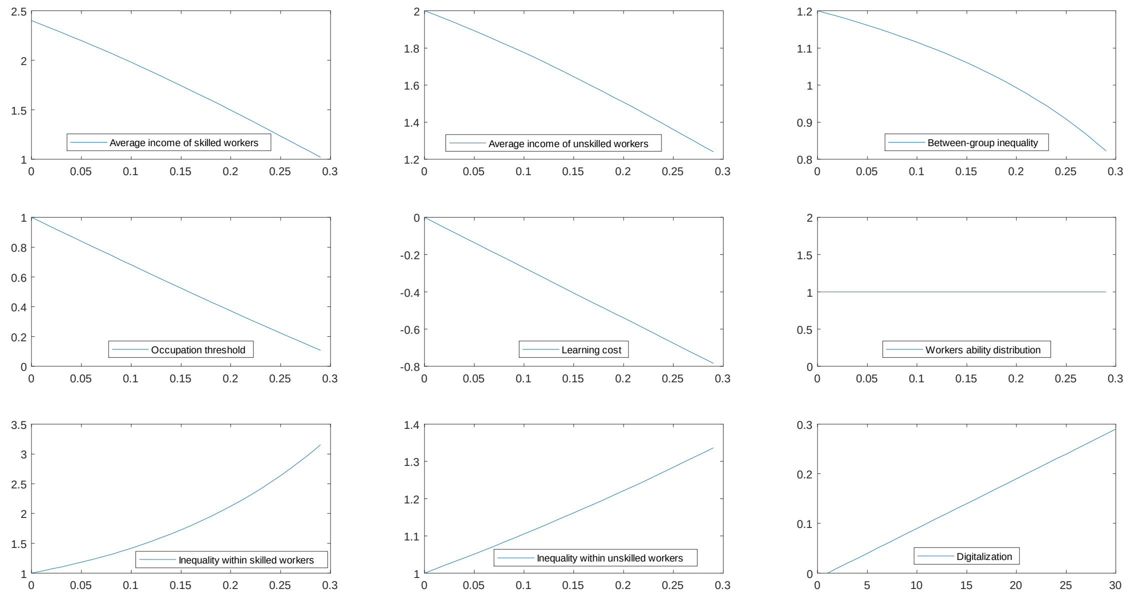

Figure 12: Model simulation with labor demand effect and endogenous learning cost ($\delta = 0.5$, $\phi = 0$, $\eta = 2.7$).

We then move to the case that learning cost effect dominates labor demand effect. This is depicted in Figure 12. This result is similar with Figure 10 except that the within-group wage income gap of skilled workers remain positive and monotonic. In addition, the extent of within-group inequality is much larger in magnitude.

5.5 Simulation with all three effects balanced

A final simulation considers an interesting case that all three effects (one labor demand channel + two labor supply channels) are switched on but their forces are still balanced for between-group inequality. Figure 13 shows this case.

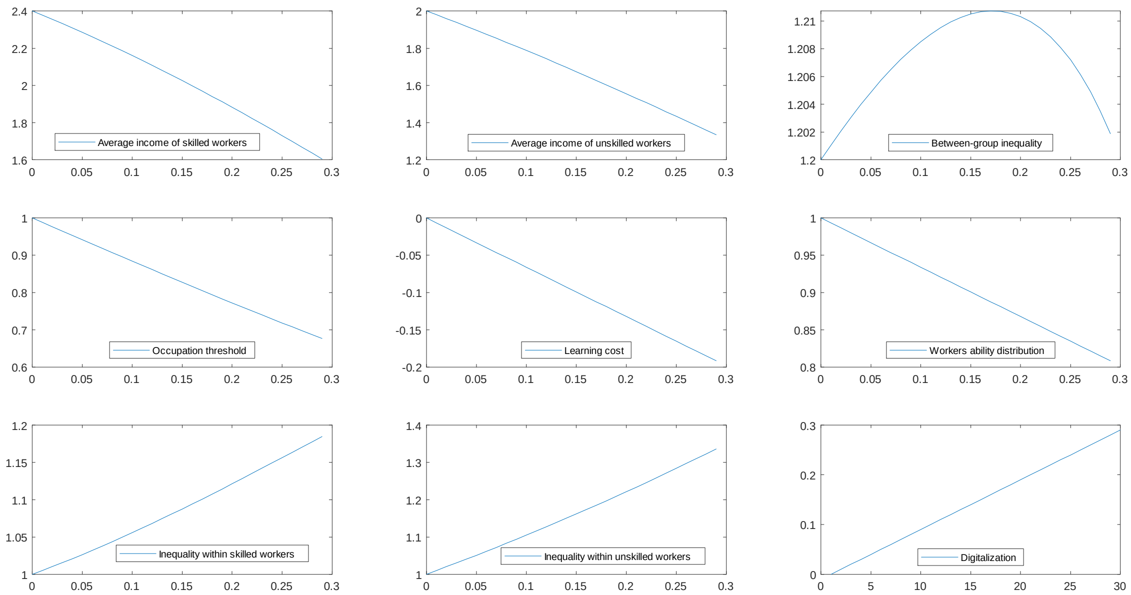

Figure 13: Model simulation with all three effects active and balanced ($\delta = 0.5$, $\phi = 0.66$, $\eta = 0.66$).

It is seen from the simulation of Figure 13 that now we only need smaller labor supply effects ($\phi = \eta = 0.66$) to offset the traditional labor demand effect. This is not surprising given that both creative destruction and learning cost channels puts downward pressure on between-group inequality (see Proposition 3 b) and c)). For with-in group inequality, it turns out to be increasing and monotonic. However, the degree of within-group inequality is much smaller than previous cases. This is because the worker ability channel offsets the upward pressure caused by learning cost channel and labor demand channel.

6 Concluding Remarks

In light of the rising income inequality across the globe since 1980s, researchers have made various explanations, many of which have acknowledged the role of the adoption of new technologies in widening wage gaps between groups. We revisit this issue by proposing a task-based growth model with digitalization and labor choice. We propose two mechanisms that can account for hump-shaped relationship between wage income inequality and digitalization. One is the worker ability channel which allows for creative destruction effect of digitalization on worker ability. The latter shifts the worker ability to the left when digitalization happens but does not affect the threshold of skilled labor. It reduces between-group wage income inequality, while decreases (increases) the within-group wage income inequality of the skilled (unskilled) workers. The other is the learning cost channel which states that workers benefit from developments of digital economy when infrastructures are more complete and platforms reduce substantially the cost of learning. It lowers the threshold of job choice and indirectly reduces between-group wage income inequality, while increases (decreases) the within-group wage income inequality of the skilled (unskilled) workers. The overall impacts of digitalization on wage income inequality then depend on the strengths of these two labor supply channels relative to the traditional labor demand channels. In particular, an inverted U-shape can show up when the three forces just offset each other.

After the theoretical analysis, a new set of empirical evidence in the Chinese economy in terms of the relationship between digitalization and wage income inequality is provided subsequently. The result shows the existence of a "financial Kuznets Curve" highlighting the importance of nonlinearity of this relationship and thus supports the main predictions of the theoretical model. The theoretical predictions are then verified in numerical simulations. Under reasonable parameterization, the model can replicate the financial Kuznets curve found in data.

The two mechanisms we propose above are novel. The worker ability channel borrows from the recent literature on automation where the creation of new tasks is positively affected by common technology progress but negatively affected by automation. The innovation of our worker ability channel is that it causes endogenous change of worker ability distribution which also causes reduces wage income inequality. The learning cost channel, on the other hand, is analogous to the positive externality effect of data economy. It captures the important insight that digital technologies differ from traditional common technology in that it may alter the occupational choice of workers. The fundamental innovativeness of these two mechanisms is that it demonstrates that digitalization not only affects labor demand by biasing production cost of firms, but also changes labor supply of workers and occupational choice which ultimately determines the dynamics of wage gaps for different groups of workers.

Our paper has important policy implications. The nonlinear relationship between wage income inequality and digitalization indicates that new technology may have both good and bad sides. Understanding both sides at different stages of economic development is critical before drawing conclusions and making policy prescriptions. Our study does support government policy that targets workers with relatively low abilities as they are less likely to be benefited from digitalization as indicated by traditional labor demand effect and the new learning cost effect. Moreover, our study also points out several directions for future research. The two mechanisms may relate to more micro-foundations in finance-related literature and the literature of knowledge which researchers should further explore. Also, our research has implications for studying other forms of new technologies, opening up new perspectives on the economic consequences of them. These leave for future research.

7 References

References

[1] Acemoglu, D. (1998). "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *The Quarterly Journal of Economics*, Oxford University Press, vol. 113(4), pages 1055-1089.

[2] Acemoglu, D. and Autor, D. (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings," Handbook of Labor Economics, in: O. Ashenfelter & D. Card (ed.), *Handbook of Labor Economics*, edition 1, volume 4, chapter 12, pages 1043-1171, Elsevier.

[3] Acemoglu, D., Makhdoumi, A., Malekian, A. and Ozdaglar, A. (2022). "Too much data: Prices and inefficiencies in data markets." *American Economic Journal: Microeconomics*, 14(4), pp.218-256.

[4] Acemoglu, D. and Restrepo, P. (2018). "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, American Economic Association, vol. 108(6), pages 1488-1542, June.

[5] Acemoglu, D. and Restrepo, P. (2020). "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, University of Chicago Press, vol. 128(6), pages 2188-2244.

[6] Acemoglu D. and Restrepo P. (2022). "Tasks, Automation, and the Rise in U.S. Wage Inequality," *Econometrica*, Econometric Society, vol. 90(5), pages 1973-2016, September.

[7] Aghion, P. (2002) "Schumpeterian Growth Theory and the Dynamics of Income Inequality." *Econometrica*, Vol. 70, No. 3 (May, 2002), 855–882.

[8] Aghion, P., Jones, B. F. and Jones, C. I. (2018). "Artificial Intelligence and Economic Growth." *NBER Chapters*, National Bureau of Economic Research, Inc.

[9] Baron, R.M. and Kenny, D. (1986). "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology*, 51, 1173-1182.

[10] Bergemann, D., Bonatti, A. and Gan, T. (2022). "The economics of social data." *The RAND Journal of Economics*, 53: 263-296.

[11] Bohm, S., Grossmann, V. and Steger, T. M. (2015). "Does expansion of higher education lead to trickle-down growth?," *Journal of Public Economics*, Elsevier, vol. 132(C), pages 79-94.

[12] Borghans, Lex and ter Weel, B. (2007). "The diffusion of computers and the distribution of wages," *European Economic Review*, Elsevier, vol. 51(3), pages 715-748, April.

[13] Brynjolfsson, E., Li, D. and Raymond, L. R. (2023). "Generative AI at Work," *NBER Working Papers* 31161, National Bureau of Economic Research, Inc.

[14] Castello-Climent, A. & Domenech, R. (2021). "Human capital and income inequality revisited," *Education Economics,* Taylor & Francis Journals, vol. 29(2), pages 194-212, March.

[15] Cong, L. W., Wei, W., Xie, D. and Zhang, L. (2022). "Endogenous growth under multiple uses of data," *Journal of Economic Dynamics and Control*, Elsevier, vol. 141(C).

[16] Cong, L. W., Xie, D. and Zhang L. (2021). "Knowledge Accumulation, Privacy, and Growth in a Data Economy." *Management Science*, 67(10):6480-6492.

[17] Che, Y. & Zhang, L. (2018). "Human Capital, Technology Adoption and Firm Performance: Impacts of China's Higher Education Expansion in the Late 1990s," *Economic Journal*, Royal Economic Society, vol. 128(614), pages 2282-2320, September.

[18] Cheng H., Jia R., Li D., & Li, H. (2019). "The rise of robots in china." *Journal of Economic Perspectives*, vol. 33(2), pages 71-88.

[19] Cortes, G. M. and Gallipoli, G. (2018). "The Costs of Occupational Mobility: An Aggregate Analysis," *Journal of the European Economic Association*, European Economic Association, vol. 16(2), pages 275-315.

[20] Dvorkin, M. and Monge-Nara A. (2019). "On the Dynamics of Occupational Choice, Human Capital Accumulation and Growth," *2019 Meeting Papers 1461*, Society for Economic Dynamics.

[21] Farboodi, M., Matray, A., Veldkamp, L. and Venkateswaran, V. (2022). "Where Has All the Data Gone?," *The Review of Financial Studies*, Volume 35, Issue 7, Pages 3101–3138.

[22] Farboodi M. and Veldkamp, L. (2021). "A Model of the Data Economy," *NBER Working Papers* 28427, National Bureau of Economic Research, Inc.

[23] Frey, C. B. (2020), *The Technology Trap: Capital, Labor, and Power in the Age of Automation*, Princeton University Press.

[24] Ge, S. and Yang, D. T. (2014). "Changes in China's Wage Structure." *Journal of the European Economic Association*, Vol. 12, No. 2, pp. 300-336

[25] Ghosh, P., Valle, B. and Zeng, Y. (2021) "FinTech Lending and Cashless Payments." *Working Paper.*

[26] Goldin, C., and Katz L. F. (2007). "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890 to 2005", *NBER Working Papers* 12984, National Bureau of Economic Research, Inc.

[27] Gomes, F., Jansson, T. and Karabulut, Y. (2022), "Do Robots Increase Wealth Dispersion?" *Working paper.*

[28] Grossman, G. M., Helpman, E., Oberfield, E. and Sampson, T. (2017). "Balanced Growth Despite Uzawa," *American Economic Review*, American Economic Association, vol. 107(4), pages 1293-1312, April.

[29] Galor, O. and Moav, O. (2000). "Ability-Biased Technological Transition, Wage Inequality, and Economic Growth," *The Quarterly Journal of Economics*, Oxford University Press, vol. 115(2), pages 469-497.

[30] He, H. and Liu, Z. (2008). "Investment-Specific Technological Change, Skill Accumulation, and Wage Inequality," *Review of Economic Dynamics*, Elsevier for the Society for Economic Dynamics, vol. 11(2), pages 314-334, April.

[31] Hemous, D. and Olsen, M. (2022). "The Rise of the Machines: Automation, Horizontal Innovation, and Income Inequality," *American Economic Journal: Macroeconomics*, American Economic Association, vol. 14(1), pages 179-223, January.

[32] Ichihashi, S. (2021). "The economics of data externalities," *Journal of Economic Theory*, Elsevier, vol. 196(C).

[33] Jones, C. I. and Tonetti, C. (2019). "Nonrivalry and the Economics of Data," *NBER Working Papers* 26260, National Bureau of Economic Research, Inc.

[34] Kambourov, G. and Manovskii, I. (2008). "Rising Occupational And Industry Mobility In The United States: 1968-97," *International Economic Review*, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 49(1), pages 41-79, February.

[35] Kambourov, G. and & Manovskii, I. (2009). "Occupational Mobility and Wage Inequality," *Review of Economic Studies*, Oxford University Press, vol. 76(2), pages 731-759.

[36] Kasa, K. and Lei, X. (2018). "Risk, Uncertainty and the Dynamics of Inequality," *Journal of Monetary Economics*, 94, 60-78.

[37] Katz, L. F. and Murphy, K. M. (1992). "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, Oxford University Press, vol. 107(1), pages 35-78.

[38] Korinek, Anton and Megan Juelfs (2022), "Preparing for the (non-existent?) future of work", NBER working paper, No. 30172

[39] Kuhn, P. and Mansour, H. (2014). "Is Internet Job Search Still Ineffective?," *Economic Journal,* Royal Economic Society, vol. 124(581), pages 1213-1233, December.

[40] Kuznets, S. (1955). "Economic Growth andIncomeInequality." *American Economic Review*, 45(1), pp.1-28.

[41] Liu, Z., Sockin, M. and Xiong, W. (2023). "Data Privacy and Algorithmic Inequality," *NBER Working Papers* 31250

[42] Messina, J. and Silva, J. (2019). "Twenty Years of Wage Inequality in Latin America," *IDB Publications (Working Papers)* 9734, Inter-American Development Bank.

[43] Milanovic, B. (2016). *Global Inequality: A New Approach for the Age of Globalization*, Harvard University Press, Belknap Press

[44] Piketty, T. (2014). *Capital in the Twenty-first Century*, Translated by A. Goldhammer, Harvard University Press.

[45] Restuccia, D. and Vandenbroucke, G. (2013). "The Evolution Of Education: A Macroeconomic Analysis," *International Economic Review*, vol. 54(3), pages 915-936, August.

[46] Zeira, J., (1998). "Workers, Machines, And Economic Growth." *Quarterly Journal of Economics*, November 1998, 113 (4), 1091–1117.

8 Appendices

A. Proofs of Propositions

A.1 Proof of PROPOSITION 1

Existence of a BGP requires that factor shares are stable. This implies that:

$$\begin{aligned}
\frac{d\left(\alpha_{K_t}\right)}{\alpha_{K_t}} &= (1-\rho)\frac{d\left(\beta_t\right)}{\beta_t}+\rho\frac{d\left(Z_t\right)}{Z_t}+\rho\left(\frac{d\left(K_t\right)}{K_t}-\frac{d\left(Y_t\right)}{Y_t}\right) \\
&= (1-\rho)\, g_{\beta_t}+\rho g_{Z_t}+\rho\left(g_{K_t}-g_{Y_t}\right)=0
\end{aligned}$$

$$\frac{d\left(\frac{\alpha_{K_t}}{\alpha_{L_t}}\right)}{\frac{\alpha_{K_t}}{\alpha_{L_t}}}=(1-\rho)\, g_{\beta_t}+\rho\left(1-\beta_t\right)\left(g_{K_t}-g_{L_t}+g_{Z_t}\right)=0$$

where we use a g_{X_t} function to denote the growth rate of variable X_t. By further making use of $g_{L_t}=n$, $g_{k_t}=g_{K_t}-n$ and $g_{y_t}=g_{Y_t}-n$, we obtain:

$$d\left(\alpha_{K_t}\right)/\alpha_{K_t}=(1-\rho)\, g_{\beta_t}+\rho g_{Z_t}+\rho\left(g_{k_t}-g_{y_t}\right)=0 \tag{25}$$

$$d\left(\frac{\alpha_{K_t}}{\alpha_{L_t}}\right)/\left(\frac{\alpha_{K_t}}{\alpha_{L_t}}\right)=(1-\rho)\, g_{\beta_t}+\rho\left(1-\beta_t\right)\left(g_{k_t}+g_{Z_t}\right)=0 \tag{26}$$

which yield:

$$g_{y_t}=\beta_t\left(g_{k_t}+g_{Z_t}\right). \tag{27}$$

Three cases of BGP in proposition 1 can be easily implied by (27). Thus, a)-c) are proved.

Combining (25) and (26), the growth rates on BGP are given by:

$$\begin{aligned}
g_{Z_t} &= \left(\frac{\rho-1}{\rho}\right) g_{\beta_t} \\
g_{k_t} &= g_{y_t}=\left[\frac{(\rho-1)\,\beta_t}{\rho\left(1-\beta_t\right)}\right] g_{\beta_t}.
\end{aligned}$$

A.2 Proof of PROPOSITION 2

Proof of PROPOSITION 2:

Proposition a) can be directly implied from (20) and (15). It is similar with the proof in Galor and Moav (2002). Propositions b) and c) can be derived by taking first order derivative of a_t^* to μ and β_t respectively. Here we provide the Proof of c):

$$\begin{aligned}
a^{*\prime}\left(\beta_t\right) &= \frac{\left(-\delta+2\delta\beta_t\right)-\eta}{\left(1+\delta\beta_t^2\right)}-\frac{\left(1-\delta\beta_t+\delta\beta_t^2+\mu_0-\eta\beta_t\right)}{\left(1+\delta\beta_t^2\right)^2}2\delta\beta_t \\
&= \frac{\begin{array}{c}\left[-\delta+2\delta\beta_t-\eta\right]\left(1+\delta\beta_t^2\right)\\ -2\delta\beta_t\left(1-\delta\beta_t+\delta\beta_t^2+\mu_0-\eta\beta_t\right)\end{array}}{\left(1+\delta\beta_t^2\right)^2} \\
&= \frac{\begin{array}{c}\left(-\delta+2\delta\beta_t-\eta\right)+\left(-\delta+2\delta\beta_t-\eta\right)\delta\beta_t^2\\ -2\delta\beta_t\left(1-\delta\beta_t+\delta\beta_t^2+\mu_0-\eta\beta_t\right)\end{array}}{\left(1+\delta\beta_t^2\right)^2} \\
&= \frac{\begin{array}{c}-\delta+2\delta\beta_t-\eta-\delta^2\beta_t^2+2\delta^2\beta_t^3-\eta\delta\beta_t^2\\ -2\delta\beta_t+2\delta^2\beta_t^2-2\delta^2\beta_t^3-2\delta\beta_t\mu_0+2\delta\beta_t^2\eta\end{array}}{\left(1+\delta\beta_t^2\right)^2} \\
&= \frac{-\delta-\eta-\eta\delta\beta_t^2+\delta^2\beta_t^2-2\delta\beta_t\mu_0+2\delta\beta_t^2\eta}{\left(1+\delta\beta_t^2\right)^2} \\
&= \frac{-\delta\left(1-\delta\beta_t^2\right)-\eta\left(1+\delta\beta_t^2\right)-2\delta\beta_t\left(\mu_0-\eta\beta_t\right)}{\left(1+\delta\beta_t^2\right)^2} \\
&= \frac{-\delta\left(1-\delta\beta_t^2\right)-\eta\left(1-\delta\beta_t^2\right)-2\delta\beta_t\mu_0}{\left(1+\delta\beta_t^2\right)^2} \\
&= \frac{-\left(1-\delta\beta_t^2\right)\left(\delta+\eta\right)-2\delta\beta_t\mu_0}{\left(1+\delta\beta_t^2\right)^2}<0
\end{aligned}$$

We can see that as long as $\delta > 0$, $\eta > 0$, $\mu_0 \geq 0$, we have $a^{*\prime}\left(\beta_t\right) < 0$ and thus c) is proved.

While this is a straightforward result, one may argue for alternative assumptions of the function forms of learning cost. Arguably, one possible specification is a quadratic cost function for learning cost:

$$\mu\left(w_t^s h_t^i,\beta_t\right)=-\frac{\eta}{2}\widetilde{A}^{\rho}w\left(k_t,\beta_t\right)\left(\beta_t-\overline{\beta}\right)^2.$$

In this case, the relationship between digitalization and the threshold may be nonlinear:

$$\begin{aligned}
\frac{\partial a_t^*\left(\beta_t\right)}{\partial\beta_t} &= \frac{\overbrace{\left(-\delta+\delta^2\beta_t^2\right)}^{\text{monotonic(negative)}}+\overbrace{\frac{\eta}{2}\left(\beta_t-\overline{\beta}\right)^2}^{\text{U-shape}}-\overbrace{\eta\left(\beta_t-\overline{\beta}\right)}^{\text{monotonic}}}{\left(1+\delta\beta_t^2\right)^2} \\
&= \frac{\overbrace{\left(-\delta+\delta^2\beta_t^2\right)}^{\text{monotonic(negative)}}+\overbrace{\eta\left(\beta_t-\overline{\beta}\right)\left(\frac{\beta_t-\overline{\beta}}{2}-1\right)}^{\text{U-shape}}}{\left(1+\delta\beta_t^2\right)^2}.
\end{aligned}$$

A.3 Proof of PROPOSITION 3

PROOF of PROPOSITION 3:

Consider first the case that there is endogenous adjustment of worker ability ($\phi > 0$) but without learning cost ($\eta = 0$). The first derivative of $\sigma_t^{\frac{S}{U}}$ to A_t is derived as follows.

$$\sigma_t^{\frac{S}{U}} = \frac{\widetilde{I}_t^S}{\widetilde{I}_t^U} = \frac{a^*\left(\beta_t\right) + A\left(\beta_t\right) - \left(2 - A\left(\beta_t\right) - a^*\left(\beta_t\right)\right)\beta_t}{\left(1 - \delta\beta_t\right)\left[2 - \left(3 - A\left(\beta_t\right) - a^*\left(\beta_t\right)\right)\beta_t\right]}.$$

$$\begin{aligned} a^*\left(\beta_t\right) &= \frac{\left(1 - \delta\beta_t + \delta\beta_t^2\right) + \left(\mu_0 - \eta\beta_t\right)}{\left(1 + \delta\beta_t^2\right)} \\ A\left(\beta_t\right) &= \widetilde{A} - \phi\beta_t \end{aligned}$$

$$\begin{aligned} a^{*\prime}\left(\beta_t\right) &= \frac{-\left(1 - \delta\beta_t^2\right)\left(\delta + \eta\right) - 2\delta\beta_t\mu_0}{\left(1 + \delta\beta_t^2\right)^2} = a^{*\prime}\left(\beta_t; \delta, \eta, \mu_0\right) < 0 \\ A'\left(\beta_t\right) &= -\phi < 0. \end{aligned}$$

$$
\begin{aligned}
\frac{\partial \sigma_t^{\frac{S}{U}}}{\partial A_t} &= \frac{1}{(1-\delta\beta_t)} \left\{ \begin{array}{c} \frac{1+\beta_t}{[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t]} \\ -\frac{[a^*(\beta_t)+A(\beta_t)-(2-A(\beta_t)-a^*(\beta_t))\beta_t]\beta_t}{\{[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t]\}^2} \end{array} \right\} \\
&= \frac{\begin{array}{c} (1+\beta_t)\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right] \\ -\left[a^*\left(\beta_t\right)+A\left(\beta_t\right)-\left(2-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\beta_t \end{array}}{(1-\delta\beta_t)\left\{\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} \\
&= \frac{\begin{array}{c} \left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right] \\ +\beta_t A'\left(\beta_t\right)\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right] \\ -\left[a^*\left(\beta_t\right)+A\left(\beta_t\right)-\left(2-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\beta_t \end{array}}{(1-\delta\beta_t)\left\{\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} \\
&= \frac{\begin{array}{c} \left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right] \\ +\beta_t \left\{ \begin{array}{c} 2-\left[3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right]\beta_t \\ -a^*\left(\beta_t\right)-A\left(\beta_t\right)+\left[2-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right]\beta_t \end{array} \right\} \end{array}}{(1-\delta\beta_t)\left\{\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} \\
&= \frac{\begin{array}{c} \left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right] \\ +\beta_t\left\{2-\beta_t-a^*\left(\beta_t\right)-A\left(\beta_t\right)\right\} \end{array}}{(1-\delta\beta_t)\left\{\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} \\
&= \frac{\left\{ \begin{array}{c} 2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t \\ +\beta_t\left[2-\beta_t-a^*\left(\beta_t\right)-A\left(\beta_t\right)\right] \end{array} \right\}}{(1-\delta\beta_t)\left\{\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} \\
&= \frac{\left\{2-\left(3-a^*\left(\beta_t\right)\right)\beta_t+\beta_t\left[2-\beta_t-a^*\left(\beta_t\right)\right]\right\}}{(1-\delta\beta_t)\left\{\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} \\
&= \frac{\left\{2-3\beta_t+a^*\left(\beta_t\right)\beta_t+2\beta_t-\beta_t^2-a^*\left(\beta_t\right)\beta_t\right\}}{(1-\delta\beta_t)\left\{\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} \\
&= \frac{\overbrace{\left\{2-\beta_t-\beta_t^2\right\}}^{\text{positive}}}{(1-\delta\beta_t)\left\{\left[2-\left(3-A\left(\beta_t\right)-a^*\left(\beta_t\right)\right)\beta_t\right]\right\}^2} > 0.
\end{aligned}
$$

It can be deduced that, given $A'\left(\beta_t\right) = -\phi < 0$, the impact of digitalization β_t on $\sigma_t^{\frac{S}{U}}$ will be negative. That is, endogenizing worker ability through creative destruction effect produces smaller wage income gap compared with a situation without such labor supply effect.

Next, consider the case that there is positive learning cost ($\eta > 0$) but no endogenous adjustment of worker ability ($\phi = 0$). In this case, only learning cost channel is in play. However, since the learning cost $\mu\left(w_t, \beta_t\right)$ only appears in the threshold a_t^*, we can easily check the relationship between $\sigma_t^{\frac{S}{U}}$ and learning cost $\mu\left(w_t, \beta_t\right)$ by examining the first derivative of $\sigma_t^{\frac{S}{U}}$ to a_t^*. It is easily verified that it is identical to the first derivative of $\sigma_t^{\frac{S}{U}}$

to A_t above. Since

$$a^{*\prime}(\beta_t) = \frac{-\left(1-\delta\beta_t^2\right)(\delta+\eta) - 2\delta\beta_t\mu_0}{\left(1+\delta\beta_t^2\right)^2} = a^{*\prime}(\beta_t;\delta,\eta,\mu_0) < 0$$

It is clear that introducing learning cost effect also helps to reduce wage income inequality.

Finally, we derive the first derivative of $\sigma_t^{\frac{S}{U}}$ to β_t to analyze the combined overall effects of the three mechanisms on wage income inequality.

$$\begin{aligned}
\frac{\partial\sigma_t^{\frac{S}{U}}}{\partial\beta_t} &= \delta\frac{a^*(\beta_t)+A(\beta_t)-(2-A(\beta_t)-a^*(\beta_t))\beta_t}{(1-\delta\beta_t)^2\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]} \\
&\quad+\frac{(1+\beta_t)\left(a^{*\prime}(\beta_t)+A'(\beta_t)\right)+(A(\beta_t)+a^*(\beta_t))-2}{(1-\delta\beta_t)\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]} \\
&\quad-\frac{\begin{bmatrix} a^*(\beta_t)+A(\beta_t) \\ -(2-A(\beta_t)-a^*(\beta_t))\beta_t \end{bmatrix}\left\{\begin{matrix} \left[a^{*\prime}(\beta_t)+A'(\beta_t)\right]\beta_t \\ -\left[3-(A(\beta_t)+a^*(\beta_t))\right] \end{matrix}\right\}}{(1-\delta\beta_t)\left\{\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]\right\}^2} \\
&= \frac{\delta\begin{bmatrix} a^*(\beta_t)+A(\beta_t) \\ -(2-A(\beta_t)-a^*(\beta_t))\beta_t \end{bmatrix}\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]}{(1-\delta\beta_t)^2\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]^2} \\
&\quad+\frac{\begin{bmatrix} (1+\beta_t)\left(a^{*\prime}(\beta_t)+A'(\beta_t)\right) \\ +(A(\beta_t)+a^*(\beta_t))-2 \end{bmatrix}(1-\delta\beta_t)\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]}{(1-\delta\beta_t)^2\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]^2} \\
&\quad-\frac{(1-\delta\beta_t)\begin{bmatrix} a^*(\beta_t)+A(\beta_t) \\ -(2-A(\beta_t)-a^*(\beta_t))\beta_t \end{bmatrix}\left\{\begin{matrix} \left[a^{*\prime}(\beta_t)+A'(\beta_t)\right]\beta_t \\ -\left[3-(A(\beta_t)+a^*(\beta_t))\right] \end{matrix}\right\}}{(1-\delta\beta_t)^2\left\{\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]\right\}^2}
\end{aligned}$$

$$\begin{aligned}
Num_t &= \delta\begin{bmatrix} a^*(\beta_t)+A(\beta_t) \\ -(2-A(\beta_t)-a^*(\beta_t))\beta_t \end{bmatrix}\begin{bmatrix} 2- \\ (3-A(\beta_t)-a^*(\beta_t))\beta_t \end{bmatrix} \\
&\quad+\begin{bmatrix} (1+\beta_t)\left(a^{*\prime}(\beta_t)+A'(\beta_t)\right) \\ +(A(\beta_t)+a^*(\beta_t))-2 \end{bmatrix}(1-\delta\beta_t)\begin{bmatrix} 2 \\ -(3-A(\beta_t)-a^*(\beta_t))\beta_t \end{bmatrix} \\
&\quad-(1-\delta\beta_t)\begin{bmatrix} a^*(\beta_t)+A(\beta_t) \\ -(2-A(\beta_t)-a^*(\beta_t))\beta_t \end{bmatrix}\left\{\begin{matrix} \left[a^{*\prime}(\beta_t)+A'(\beta_t)\right]\beta_t \\ -\left[3-(A(\beta_t)+a^*(\beta_t))\right] \end{matrix}\right\} \\
&= Num\left(\beta_t;\delta,\eta;\widetilde{A},\mu_0\right). \\
\frac{\partial\sigma_t^{\frac{S}{U}}}{\partial\beta_t} &= \frac{Num\left(\beta_t;\delta,\eta;\widetilde{A},\mu_0\right)}{(1-\delta\beta_t)^2\left[2-(3-A(\beta_t)-a^*(\beta_t))\beta_t\right]^2} = f\left(\beta_t;\delta,\eta;\widetilde{A},\mu_0\right).
\end{aligned}$$

It can be seen that the derivative $\frac{\partial\sigma_t^{\frac{S}{U}}}{\partial\beta_t}$ is rather tedious, and we will examine this numerically in the Simulation section. However, the principles are very clear: the ultimate relationship between digitalization and wage income inequality depends on the relative strengths of the labor demand effect of SBTC (enlarging wage gap) and the two labor supply effects (shrinking wage gap). Those effects are governed ultimately by three key parameters $\{\delta, \phi, \eta\}$. In particular, there exists a combination $\{\delta^*, \phi^*, \eta^*\}$ that making $\frac{\partial \sigma_t^{\frac{S}{U}}}{\partial \beta_t} = 0$ which delivers an inverted U-shape with $\beta_t^* \{\delta^*, \phi^*, \eta^*\}$ as the inflection point.

A.4 Proof of PROPOSITION 4

The proof follows the same strategy with the proof of PROPOSITION 3, thus the presentation is skipped.

the *within-group* of skilled workers is computed as:

$$\sigma_t^S = \frac{i^S (A_t)}{i^S (a_t^*)} = \frac{A_t - (1 - A_t) \beta_t}{a_t^* - (1 - a_t^*) \beta_t}. \tag{28}$$

Similarly, *within-group* wage inequality of unskilled workers is computed as:

$$\sigma_t^U = \frac{i^U (a_t^*)}{i^U (A_t - 1)} = \frac{1 - (1 - a_t^*) \beta_t}{1 - (2 - A_t) \beta_t}. \tag{29}$$

$$a_t^* = \frac{\left(1 - \delta\beta_t + \delta\beta_t^2\right) + (\mu_0 - \eta\beta_t)}{\left(1 + \delta\beta_t^2\right)} \tag{30}$$

$$A_t = \widetilde{A} - \phi\beta_t. \tag{31}$$

$$\mu\left(w_t^s h_t^i, \beta_t\right) = w_t^s h_t^i (\mu_0 - \eta\beta_t). \tag{32}$$

First of all, because $\frac{\partial a_t^*}{\partial A_t} = 0$, we have

$$\begin{aligned} \frac{\partial \sigma_t^S}{\partial A_t} &= \frac{1 + \beta_t}{a_t^* - (1 - a_t^*) \beta_t} > 0 \\ \frac{\partial \sigma_t^U}{\partial A_t} &= -\frac{1 - (1 - a_t^*) \beta_t}{\left[1 - (2 - A_t) \beta_t\right]^2} \beta_t < 0 \end{aligned}$$

Since $A'(\beta_t) = -\phi < 0$, the creative destruction channel itself implies that digitalization reduces worker ability, and the same time reduces within group inequality of skilled workers while raises within group inequality of unskilled workers. Thus, b) is proved.

Second, because $\frac{\partial A_t}{\partial a_t^*} = 0$, we have

$$\begin{aligned} \frac{\partial \sigma_t^S}{\partial a_t^*} &= -\frac{A_t - (1 - A_t) \beta_t}{\left[a_t^* - (1 - a_t^*) \beta_t\right]^2} (1 + \beta_t) < 0 \\ \frac{\partial \sigma_t^U}{\partial a_t^*} &= \frac{\beta_t}{1 - (2 - A_t) \beta_t} > 0 \end{aligned}$$

Since $a^{*\prime}(\beta_t) = -\phi < 0$, the learning cost channel itself implies that digitalization reduces job threshold, and the same time raises within group inequality of skilled workers while reduces within group inequality of unskilled workers. Thus, c) is proved.

To see overall effects on within group inequality, we derive

$$\begin{aligned}
\frac{\partial \sigma_t^S}{\partial \beta_t} &= \frac{A'(\beta_t) - (1 - A_t) + \beta_t A'(\beta_t)}{a_t^*(\beta_t) - (1 - a_t^*(\beta_t))\beta_t} \\
&\quad - \frac{[A_t - (1 - A_t)\beta_t]\left[a^{*\prime}(\beta_t) - (1 - a_t^*(\beta_t)) + \beta_t a^{*\prime}(\beta_t)\right]}{[a_t^*(\beta_t) - (1 - a_t^*(\beta_t))\beta_t]^2} \\
&= \frac{\begin{array}{c}[(1+\beta_t)A'(\beta_t) - (1 - A_t)][(1+\beta_t)a_t^*(\beta_t) - \beta_t] \\ -[(1+\beta_t)a^{*\prime}(\beta_t) - (1 - a_t^*(\beta_t))][(1+\beta_t)A_t - \beta_t]\end{array}}{[a_t^*(\beta_t) - (1 - a_t^*(\beta_t))\beta_t]^2} \\
&= \frac{Num_t^{\sigma_t^S}}{[a_t^*(\beta_t) - (1 - a_t^*(\beta_t))\beta_t]^2} \\
&= f\left(\beta_t; \delta, \eta; \widetilde{A}, \mu_0\right).
\end{aligned}$$

Since both $A'(\beta_t)$ and $a^{*\prime}(\beta_t)$ are negative, the sign of $Num_t^{\sigma_t^S}$ and $\frac{\partial \sigma_t^S}{\partial \beta_t}$ are uncertain. Given the form of $Num_t^{\sigma_t^S}$, and it is deduced that the sign of $\frac{\partial \sigma_t^S}{\partial \beta_t}$ depends on the difference $[A'(\beta_t) - a^{*\prime}(\beta_t)]$ which is basically a function of $\left(\beta_t; \delta, \eta; \widetilde{A}, \mu_0\right)$ and ϕ respectively. In other words, the relative dominance of the three key effects determines whether within group inequality shrinks or not.

Finally,

$$\sigma_t^U = \frac{i^U(a_t^*)}{i^U(A_t - 1)} = \frac{1 - (1 - a_t^*(\beta_t))\beta_t}{1 - (2 - A_t)\beta_t}$$

$$\begin{aligned}
\frac{\partial \sigma_t^U}{\partial \beta_t} &= \frac{a_t^{*\prime}(\beta_t)\beta_t - (1 - a_t^*(\beta_t))}{1 - (2 - A_t(\beta_t))\beta_t} \\
&\quad - \frac{[A_t'(\beta_t)\beta_t - (2 - A_t(\beta_t))][1 - (1 - a_t^*(\beta_t))\beta_t]}{[1 - (2 - A_t(\beta_t))\beta_t]^2} \\
&= \frac{\begin{array}{c}[a_t^{*\prime}(\beta_t)\beta_t - (1 - a_t^*(\beta_t))][1 - (2 - A_t(\beta_t))\beta_t] \\ -[A_t'(\beta_t)\beta_t - (2 - A_t(\beta_t))][1 - (1 - a_t^*(\beta_t))\beta_t]\end{array}}{[1 - (2 - A_t(\beta_t))\beta_t]^2} \\
&= \frac{Num_t^{\sigma_t^U}}{[1 - (2 - A_t(\beta_t))\beta_t]^2},
\end{aligned}$$

$$
\begin{aligned}
& Num_t^{\sigma_t^U} \\
= \; & \left[a_t^{*\prime}(\beta_t)\beta_t - (1 - a_t^*(\beta_t))\right]\left[1 - (2 - A_t(\beta_t))\beta_t\right] - \left[A_t'(\beta_t)\beta_t - (2 - A_t(\beta_t))\right]\left[1 - (1 - a_t^*(\beta_t))\beta_t\right] \\
= \; & \left[a_t^{*\prime}(\beta_t)\beta_t - (1 - a_t^*(\beta_t))\right] - \left[a_t^{*\prime}(\beta_t)\beta_t - (1 - a_t^*(\beta_t))\right](2 - A_t(\beta_t))\beta_t \\
& - \left[A_t'(\beta_t)\beta_t - (2 - A_t(\beta_t))\right] + \left[A_t'(\beta_t)\beta_t - (2 - A_t(\beta_t))\right](1 - a_t^*(\beta_t))\beta_t \\
= \; & \left[a_t^{*\prime}(\beta_t)\beta_t - A_t'(\beta_t)\beta_t\right] + \left[(1 - a_t^*(\beta_t))\beta_t^2 A_t'(\beta_t) - (2 - A_t(\beta_t))\beta_t^2 a_t^{*\prime}(\beta_t)\right] \\
& + \left[2 - A_t(\beta_t) - 1 + a_t^*(\beta_t)\right] + \left[(1 - a_t^*(\beta_t))(2 - A_t(\beta_t))\beta_t - (2 - A_t(\beta_t))(1 - a_t^*(\beta_t))\beta_t\right] \\
= \; & \left[a_t^{*\prime}(\beta_t)\beta_t - A_t'(\beta_t)\beta_t\right] + \left[(1 - a_t^*(\beta_t))\beta_t^2 A_t'(\beta_t) - (2 - A_t(\beta_t))\beta_t^2 a_t^{*\prime}(\beta_t)\right] \\
& + \left[1 - A_t(\beta_t) + a_t^*(\beta_t)\right] + \beta_t(2 - A_t(\beta_t))\left[(1 - a_t^*(\beta_t)) - (1 - a_t^*(\beta_t))\right] \\
= \; & \left[a_t^{*\prime}(\beta_t)\beta_t - A_t'(\beta_t)\beta_t\right] + \left[(1 - a_t^*(\beta_t))\beta_t^2 A_t'(\beta_t) - (2 - A_t(\beta_t))\beta_t^2 a_t^{*\prime}(\beta_t)\right] \\
& + \left[1 - A_t(\beta_t) + a_t^*(\beta_t)\right]
\end{aligned}
$$

Again, the overall effect of digitalization on wthin-group inequality of unskilled workers depend on the relative size of $A'(\beta_t)$ and $a^{*\prime}(\beta_t)$ and also β_t. We explored the property of $\frac{\partial \sigma_t^U}{\partial \beta_t}$ numerically in Simulation section.