Automation, Human Task Innovation, and Labor Share: Unveiling the Role of Elasticity of Substitution[*]

Seungjin Baek,[†] Deokjae Jeong[‡][§]

Abstract

This paper investigates the elements contributing to the change in labor share, with a specific focus on the roles of 'automation' and 'innovation in human tasks.' We construct a general equilibrium model that distinctly incorporates both robot and non-robot capital to derive an econometric specification. Using task data from O*NET and employing the most recently developed sentence embedding tools to match tasks and patents, we construct a novel 'innovation in human tasks' variable for multiple countries. This allows us to empirically evaluate the impact of innovation in human tasks on labor share across countries for the first time in the literature. Our accounting analysis suggests that the positive influence of human task innovation outweighs the adverse effects of automation in most of countries we study. From our regression analysis, we estimate the elasticity of substitution between labor and non-robot capital to be less than one, while the elasticity of substitution between tasks is greater than one. With these estimates, we elucidate the direct and indirect effects of automation and innovation in human tasks on labor share.

Keywords: innovation in human tasks, automation, labor share, elasticity of substitution

JEL Codes: D24, E24, E25, J23, O33, O57

[*]We extend our heartfelt thanks to Giovanni Peri for his ongoing guidance and invaluable support. We are also deeply grateful to Òscar Jordà, Athanasios Geromichalos, Colin Cameron, Alan Taylor, Takuya Ura, Katheryn Russ, Marianne Bitler, Monica Singhal, Jenna Stearns, and Mark Siegler for their invaluable advice and insights throughout the course of this project. I also thank participants at the WEAI 99th Annual Conference, the American Economic Association Annual Meeting in 2024, the Annual All-California Labor Economics Conference, Korea-America Economic Association Job Market Conference, UC Davis Macro Lunchtime seminar, and UC Davis Applied Micro Lunchtime seminar for their helpful comments and discussion. All errors are our own.

[*]Replication data and code and the most recent version of paper:
https://github.com/jayjeo/public/blob/main/Laborshare/readme.md

[†]University of California, Davis; sjsbaek@ucdavis.edu; https://sites.google.com/view/sjbaek-com

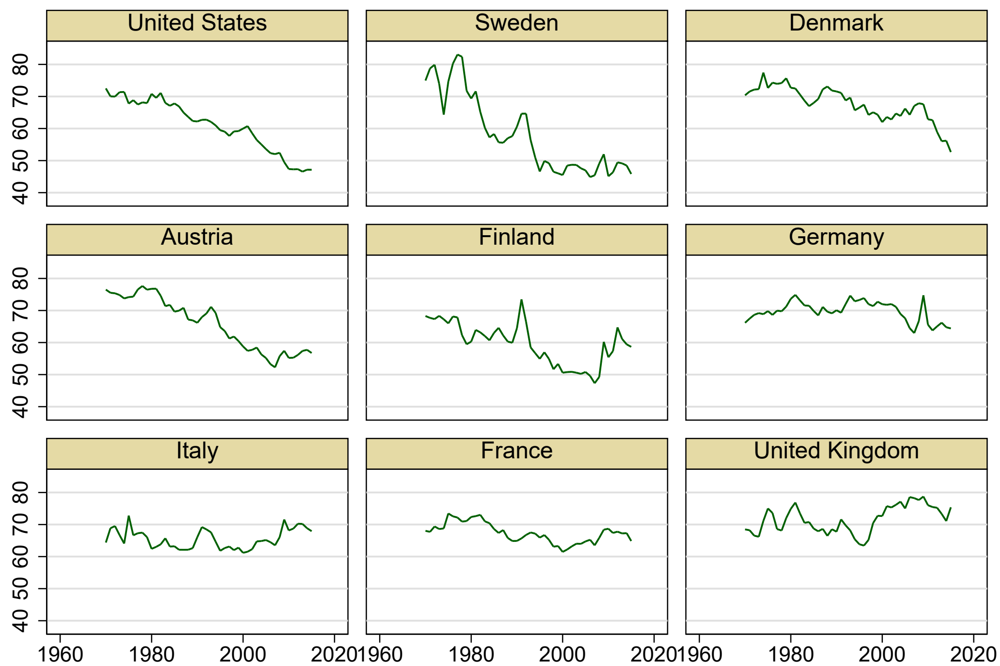
[‡]SSK Inclusive Economic Policy Research Team, South Korea; ubuzuz@gmail.com; jayjeo.com

[§]Corresponding author

1 Introduction

Karabarbounis and Neiman (2014) and Autor et al. (2020) have noted that the global labor share has followed a declining trend since the early 1980s, with an average decrease of about five percentage points. Figure 1, based on data compiled by Gutiérrez and Piton (2020), compares the labor shares in the manufacturing sector between the USA and the eight EU nations that we studied. Note that this study covers ten countries, including Portugal, which is omitted in this figure. While the USA, Sweden, Denmark, and Austria have witnessed significant declines, other countries report comparatively slight decreases. This discrepancy indicates that global labor share trends exhibit considerable heterogeneity, further underscoring our aim to investigate variations across countries and sectors to better understand this decline.[1]

Figure 1: Labor shares



Although the precise cause of this decline is still a topic of debate, advancements in automation emerge as a possible key driver. The urgency of addressing the diminishing labor share intensifies with the accelerated growth in automation and artificial intelligence technologies. For instance, Boston Dynamics has unveiled Atlas, a hu-

[1]In this context, our study aligns with Graetz and Michaels (2018), which assesses seventeen EU countries, although their focus is predominantly on productivity growth rather than the decrease in labor share.

manoid robot with impressive speed and capabilities.[2] The recent debut of Chat-GPT 4, which astoundingly achieved a 10% ranking in the United States bar exam, further underscores the rapid evolution of AI systems.[3]

The influence of automation on labor share remains a prominent topic in active research. Several studies such as those by Acemoglu and Restrepo (2020), Acemoglu et al. (2020), Dauth et al. (2021), and Martinez (2018) suggest that automation reduces labor share. In contrast, findings from research like De Vries et al. (2020) and Gregory et al. (2016) propose that automation amplifies labor share. Moreover, studies by Humlum (2019) and Hubmer and Restrepo (2021) explore the diverse impacts of automation on various population groups and industry sectors.

Yet, another factor potentially promoting labor share is the 'innovation in human tasks' —innovative tasks beyond the capabilities of robots. Autor (2015) contends that the sustained relevance of human labor in the future will largely depend on the pace at which 'innovation in human tasks' outstrips the advancement of automation. To the best of our knowledge, Autor et al. (2024) is the only study that empirically measures the innovation in human tasks.[4] While our measurement for 'innovation in human tasks' uses a different approach and methodology, we also utilize patent information, as Autor et al. (2024) does. This variable incorporates country variations derived from the US Census and EU-LFS data, enabling a global investigation of this effect on labor share. Our findings indicate that innovation in human tasks serves as an effective counterbalance to the negative effects of automation on labor share across the countries studied.

Automation and innovation in human tasks are not the only factors contributing to changes in labor share. In literature, many other reasons have been meticulously examined, especially using causality techniques. However, fewer studies attempt to measure multiple reasons within a unified framework (Bergholt et al., 2022).[5] Grossman and Oberfield (2022) highlighted the importance of utilizing general equilibrium analysis, stating: "Many authors present different sides of the same coin ... Even if the various mechanisms are all active, it becomes difficult to gauge what part of the effect estimated in one study has already been accounted for elsewhere." To address this challenge, we adopt a general equilibrium model, an approach that represents a contribution to the existing literature. The study most akin to ours is that of Acemoglu

[2] https://youtu.be/-e1_QhJ1EhQ

[3] https://youtu.be/EunbKbPV2C0

[4] Kogan et al. (2023)'s work is also relevant, albeit different from that of Autor et al. (2024) and our study. In each occupation, they measure the degree of exposure to automation technologies and labor-augmenting technologies. The former pertains to technologies associated with routine tasks, while the latter concerns technologies related to non-routine tasks.

[5] Bergholt et al. (2022) points out that "while a large literature has discussed each of these four explanations in isolation, an empirical analysis including all of them in the context of the same model is lacking. Our aim is to fill this gap."

and Restrepo (2022b). They too utilize a general equilibrium model, though their main focus is on wage inequality rather than the decline in labor share.

Specifically, our analysis incorporates five potential determinants within our general equilibrium model: automation, innovation in human tasks, capital price, robot price, and wages.[6] Based on the econometric specification derived from the model and the regression results, we estimate that the elasticity of substitution *between labor and non-robot capital* is less than one, while the elasticity of substitution *between labor and robot capital* is estimated to be greater than one.

Based on these estimates, we clarify the mechanisms by which five potential determinants influence labor share. First, we observe that automation negatively affects labor share, while innovation in human tasks positively affects it, as the literature theoretically predicts and empirically suggests. Furthermore, we suggest that both the negative effect of automation and the positive effect of innovation in human tasks are amplified through the aggregated task price channel: First, automation and innovation in human tasks alter the composition of tasks performed by robots and those performed by labor. Second, this change in composition affects the aggregate task price. Finally, the change in the aggregate task price, in turn, affects labor share through substitution among labor, robots, and non-robot capital.

Second, the regression results show a positive association between labor price and labor share and a negative association between the price of non-robot capital and labor share. The underlying intuition stems from the gross complementarity between labor and non-robot capital. Based on these results, we add empirical evidence that the elasticity of substitution *between labor and capital* is less than one, which the majority of the literature supports, as suggested by Chirinko (2008) and Grossman and Oberfield (2022). Even though our finding on the elasticity is not consistent with Karabarbounis and Neiman (2014), our results show that the price of non-robot capital has caused labor share to decline, aligning with Karabarbounis and Neiman (2014), as our data suggest that non-robot capital prices have generally increased over the past 15 years.

Third, the regression results provide a positive but insignificant association between robot price and labor share —when robot price declines, the labor share would slightly decrease. The insignificance is attributable to the current low share of robot cost among the total costs. However, our model and emprical results anticipate that as automation becomes more prevalent in the future, lower robot prices or higher robot productivity can significantly decrease labor share. This prediction is crucially based

[6] In this context, the research by Bergholt et al. (2022) closely aligns with our study. They examine rising markups, increased worker bargaining power, a declining investment price, and escalating automation as factors contributing to the falling labor share. Although their methodology, which employs time series techniques (Structural VAR with sign restrictions) and focuses exclusively on the USA, differs from ours, their findings are in line with our results. They identify automation as a principal driver of the reduction in labor share. Interestingly, they conclude that a declining capital price does not contribute to the decrease in labor share.

on the estimation results showing that the elasticity of substitution between labor and robot capital is much larger than that between labor and non-robot capital.[7] These results contribute to literature by empricially supporting the theoretical condition that improvement in robot productivity significantly decreases labor share both in short and long-run, as discussed in Berg et al. (2018).

In the following section, we provide key definitions used in this study. In Section 3, we present our general equilibrium model, while Section 4 details the datasets we used. Section 5 conducts the regression analysis, and Section 6 performs various accountings to ascertain which mechanism predominantly explains labor share decline across different countries and industries. Section 7 offers various robustness checks to demonstrate that all our intuitions and results from the main analysis remain stable across different specifications. Finally, Section 8 provides our concluding remarks.

2 Definitions

This section provides definitions for 'robot', 'automation', and 'innovation in human tasks' that will be used throughout this paper. We adhere to the definition of a robot as specified by ISO standard 8373:2012, which describes it as an "automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes."[8] The International Federation of Robotics (IFR) also strictly adheres to this definition (Müller, 2022). We source our robot data from the IFR.

In Figure 2, Panel (a) depicts a robot. However, Panel (b) is not robot because this milling machine does not come with any type of hook-up to have it run automatically. Therefore, it is neither reprogrammable nor automatically controlled. Additionally, it cannot be considered multipurpose, as it is designed solely for milling. Also, it does not operate on three or more axes. This example underscores the narrow definition of a robot.

We define 'automation' as the enhancement of robots' capabilities, which allows them to perform tasks that were previously unachievable. This definition is consistent with the one provided by Acemoglu and Restrepo (2018) and Acemoglu and Restrepo (2019).

Lastly, we define 'innovation in human tasks' as the expansion of tasks that human-workers are expected to perform because those are beyond the capabilities of robots.

[7] This means when robot prices decrease, robots significanlty take from labor's share. However, labor cannot take a significant share from non-robot capital, even though labor is more productive due to lower robot prices.

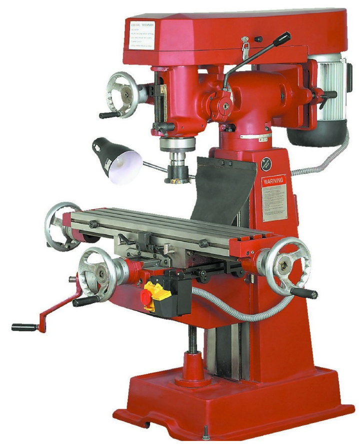
[8] Acemoglu and Restrepo (2020) also defines robots in a manner consistent with this description: "fully autonomous machines that do not need a human operator and can be programmed to perform several manual tasks ... This definition excludes other types of equipment."

[9] Vertical milling machine by harborfreight

Figure 2: Examples of Robot

(a) Robot

(b) Not robot[9]



For instance, according to ONET, the job description for Urban and Regional Planners (SOC 19-3051) expanded from 19 responsibilities in 2019 to include tasks related to statistics and data management. Previously, their responsibilities included: (1) holding public meetings with officials and scientists, (2) advising planning officials on project feasibility and cost-effectiveness, and (3) mediating community disputes. One year later, their scope of tasks widened to incorporate: (1) preparing reports using statistics, (2) developing and maintaining maps and databases, and (3) researching, compiling, analyzing, and organizing information. This serves as a prototypical example of innovation in human tasks, illustrating that individuals aspiring to become Urban and Regional Planners must now acquire skills in data handling and statistics. The integration of new tasks represents more than a mere redistribution of responsibilities from statisticians to Urban and Regional Planners. It necessitates their proficiency in statistical techniques to produce novel insights, combining their expertise in urban planning with statistical knowledge.

It is important to note that our definition of 'innovation in human tasks' includes not only new tasks in specific occupations but also the expansion of tasks that already exist in other occupations. For example, skills in data handling and statistics, previously required in other fields, are now also necessary for Urban and Regional Planners. This definition aligns with the 'upgrade in tasks' concept used by Acemoglu and Restrepo (2018) and Acemoglu and Restrepo (2019), but it contrasts with Autor et al. (2024). They define 'innovation in human tasks', or 'new tasks' in their terminology, strictly as the creation of entirely new occupations that previously did not exist.

It is important to note that the definition of 'innovation in human tasks' does not inherently guarantee increases in wages, employment, or labor share; in fact, it may lead to decreases. This is consistent with other studies (Acemoglu and Restrepo, 2018; Autor et al., 2024). Specifically, Acemoglu and Restrepo (2018) demonstrates, through a proposition, that under certain parameters, 'innovation in human tasks' indeed leads to increases in wages and employment. Similarly, Autor et al. (2024) finds empirical evidence that "employment and wage bills expand in occupations exposed to 'augmentation innovation' " (referred to as 'innovation in human tasks' in our terminology).[10]

3 Model

Acemoglu and Restrepo (2018) have offered a formal model that outlines how labor share is influenced by 'automation' and 'innovation in human tasks.' We have refined our model based on their static version. Our key contribution is the distinction we make between robots and other capital equipment, a distinction their model does not delineate. Acemoglu and Restrepo (2020) found that advancements in robotics negatively impact wages and employment. Conversely, they discovered that other forms of capital positively impact these variables. This distinction emphasizes that 'robots' and 'capital' can carry different implications for labor demand.

Our model holds advantages over existing literature, such as Berg et al. (2018) and DeCanio (2016), which also introduced robots as a separate factor from traditional capital. First, our model comprehensively incorporates factors affecting labor share, most importantly automation and innovation in human tasks, in addition to factor prices. This allows us to quantitatively analyze the extent to which each factor affects labor share across different sectors and countries. Second, our model delivers in-depth interpretations regarding the substitutability between labor, capital, and robots. From the regression equations derived from the task-based model, we gain unique insights into the degree of substitutability among factors, as well as the tasks conducted by either labor or robots.

3.1 Firms

In our model, firms face monopolistic competition, which allows them to generate positive profits. For simplicity, we assume that the production function is the same for all firms[11]. Also, for brevity, we omit the time subscript.

[10] Meanwhile, Acemoglu and Restrepo (2019) employs a strong assumption that 'reinstatement' necessarily increases the labor share. Although they did not explicitly state, we surmise that their assumption might be supported by the propositions made by Acemoglu and Restrepo (2018). Utilizing this assumption, Acemoglu and Restrepo (2019) empirically infer the reinstatement through a decomposition of labor share. For a comprehensive explanation of their methodology, see Appendix A.

[11] Introducing heterogeneity in terms of Hicks-neutral productivity does not change our analysis.

Each firm utilizes a continuum of tasks, indexed between $N-1$ and N, in addition to capital, for production. As in Acemoglu and Restrepo (2018), N increases over time due to innovation in human tasks, which can only be conducted by labor. Additionally, there is an index I that falls between $N-1$ and N. I is related to the possibility of automation and thus increases along with improvements in automation technology. Specifically, tasks below I in firm i can technically be conducted by either labor or robots, while tasks above I can only be performed by labor, as follows:

$$t_j(i) = m_j(i) + \gamma_j l_j(i) \ \text{ if } j \leq I \tag{1}$$

$$t_j(i) = \gamma_j l_j(i) \ \text{ if } j > I \tag{2}$$

, where $m_j(i)$ and $l_j(i)$ represent the number of robots and labor used for task j in firm i. γ_j represents the productivity of labor for task j. The productivity, γ_j, increases with a higher task index, j.

Tasks, $t_j(i)$, are aggregated using Constant Elasticity of Substitution (CES) aggregator, and both the aggregated tasks and capital are further combined using another CES function. Therefore, the production function is:

$$Y(i) = \left(T(i)^{\frac{\sigma-1}{\sigma}} + K(i)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{3}$$

$$T(i) = \left(\int_{N-1}^{N} t_j(i)^{\frac{\zeta-1}{\zeta}} dj\right)^{\frac{\zeta}{\zeta-1}} \tag{4}$$

, where $T(i)$ and $K(i)$ represent the number of aggregated tasks and capital used for the production of the final good i, denoted as $Y(i)$. Meanwhile, σ and ζ represent the elasticity of substitution between *aggregated tasks and non-robot capital*, and the elasticity of substitution between *tasks*, respectively.

Factor markets are assumed to be perfectly competitive. Additionally, since we focus on long-run change in labor share, it is reasonable to assume that factors are supplied elastically. For further simplicity, we assume that factors are supplied perfectly elastically at a given factor price at each period.

3.2 Labor Share

Let us move the detailed elaboration of our model to Appendix B. Based on Equations (15) to (22) presented in this appendix, the labor share is derived as follows:

$$S_L = \frac{\eta - 1}{\eta} \frac{\int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj}{P_T^{1-\zeta}} \frac{P_T^{1-\sigma}}{P_T^{1-\sigma} + R^{1-\sigma}} \tag{5}$$

$$, \text{where } P_T \equiv \left[(I-N+1)\psi^{1-\zeta} + \int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj\right]^{\frac{1}{1-\zeta}}$$

, where γ_j represents the productivity of labor for task j. The productivity, γ_j, increases with a higher task index, j. W_j, ψ, and R represent wage for labor conducting task j, robot price, and capital price, respectively.

It is worth mentioning that the term, $\frac{\eta-1}{\eta}$, is the inverse of the firm's mark-up. Since we focus on labor income as a fraction of total factor income, we denote it as S_L^f as follows:

$$S_L^f \equiv \frac{\eta}{\eta-1} S_L = \frac{\int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj}{P_T^{1-\zeta}} \frac{P_T^{1-\sigma}}{P_T^{1-\sigma} + R^{1-\sigma}} \tag{6}$$

In the next section, we discuss the datasets used in this paper and the construction of the variables.

4 Data Collection and Variable Generation

For the purpose of assessing 'innovation in human tasks' (henceforth referred to as IHT), we will use data from ONET, which offers information on the number of new tasks in the USA, measured at the occupation-year level. This data is collected directly by ONET. To analyze automation, we will use data provided by the International Federation of Robotics (IFR), which gives us the number of automated machines at the country-industry-year level.

4.1 Innovation in Human Tasks

By taking the natural log of Equation (6) and then computing the total derivative of the resulting equation with respect to the exogenous variables in the model (I, N, $\mathbb{W}$, ψ, R, and γ), we obtain Equation (9) that appears in Regression section. This equation represents our final regression equation.

To proxy dN in Equation (9), we will use IHT, which we elaborate in this subsection. The Occupational Information Network (ONET), managed and maintained by the United States Department of Labor, serves as a comprehensive database of occupational information (National Center for O*NET Development, 2023). For each Standard Occupational Classification (SOC),[12] ONET consistently updates the spectrum of tasks that

[12] SOC is an acronym for Standard Occupational Classification employed by US agencies. The ONET classification system (ONET-code) is a subclassification of the SOC system, hence, every ONET-code

workers are expected to perform. For example, in 2023, Automotive Engineers were assigned 25 responsibilities, which included the calibration of vehicle systems, control algorithms, and other software systems. When new tasks, previously nonexistent, come to light, ONET increases the number of tasks associated with the Automotive Engineering occupation.

Furthermore, ONET periodically reports 'Emerging new tasks' about once or twice annually. These tasks have recently emerged but have not been extensively studied by the ONET department; hence, these specific tasks are not included in the occupational list. We incorporate these 'Emerging new tasks' in addition to our base number of tasks provided by ONET. This process completes our generation of 'Task scores' by each occupation.[13]

The 'Task scores' vary by Standard Occupational Classification (SOC) and year. Acemoglu and Restrepo (2019) (henceforth AR) translated this information into variations by industry and year using the US Census from IPUMS (Ruggles et al., 2020), a dataset comprising individual worker data with specific occupation codes.[14] After associating the 'Task score' with each individual, an average is calculated at the industry and year level. Subsequently, we compute the 5-year growth rate of this variable, which we denote as IHT. It can also be formulated for EU countries using the EU Labor Force Survey (EU-LFS) instead of the US Census. It's crucial to recognize that the 'Task scores' from ONET are used to generate IHT for EU countries.

The European Commission has recently initiated a project akin to ONET, named 'European Skills, Competences, Qualifications, and Occupations' (ESCO). ESCO has disclosed the tasks required for workers for a single year and has yet to release a Task score. In the absence of a European equivalent of the 'Task scores', we depend on data from ONET. A foundational assumption in the creation of the EU's IHT is that the task requirements in the USA mirror similar trends in the EU. For example, if the number of tasks required for Automotive Engineers surged in the USA in 2015, it is

has a corresponding SOC. However, the ONET-code does not align perfectly with the Occupational Classification Code (OCC).

[13]Meanwhile, Acemoglu and Restrepo (2019) employs only 'Emerging new tasks' to construct the Task scores. We contend that our method of integrating both the 'base number of tasks' and 'Emerging new tasks' offers a more sophisticated approach than relying solely on Emerging new tasks, as AR does. Specifically, the 'base number of tasks' serves as a primary source of information for capturing new tasks that were nonexistent before, while 'Emerging new tasks' function as supplementary information.

[14]Our matching procedure from 'Task score' to the US Census is as follows: We use SOC as it is, instead of converting it to OCC as Acemoglu and Restrepo (2019) does. The US Census provides both SOC and OCC for occupational taxonomy, allowing us to simply use SOC to match the US Census with the 'Task score'.

Moreover, when matching 'Task score' to EU-LFS, using SOC is more advantageous than using OCC. EU-LFS uses ISCO for occupational taxonomy, and ISCO (4-digits) matches with SOC (6-digits). This granular level of crosswalk matching is made possible by the recent work of Frugoli and ESCO (2022). The excel file for the crosswalk between ISCO and SOC is in this link. This is publicly released by ONET and ESCO.

assumed that a similar trend occurred in the EU around the same period. Therefore, the variation for the EU originates from the differing composition of workers in each country, occupation, and year.

A potential criticism concerns the 'Task scores' assigned to each occupation and year. For instance, is the increment of one task by Political Science Teachers equivalent to that by Aerospace Technicians? AR's methodology treats all tasks inherently valued equally. In contrast, we apply weights to the 'Task scores.' Specifically, we match occupation descriptions with US patents descriptions granted each year through a many-to-many matching approach, and calculate the cumulative similarity scores between them. For example, consider 2018: suppose ten patents matched to Aerospace Technicians[15] yielded a total similarity score of 6.5, then an increase of one task for Aerospace Technicians in 2018 is valued at $1 \times 6.5 = 6.5$. Conversely, if a single patent was matched to Political Science Teachers[16], and it had a similarity score of 0.7, then an increase of one task for Political Science Teachers in the same year would be valued at $1 \times 0.7 = 0.7$.

We acknowledge that a precise definition of IHT encompasses the development of tasks not inherently limited to those related to patents. In other words, why should a new task devised by Political Science Teachers be valued less than one created by Aerospace Technicians? To address this issue, we will also provide regression results and accounting analysis from an unweighted version in the Robustness Check section.

Let us provide further details about our patent matching method. Unlike other studies, we utilized the most recently developed text-to-vector embedding software. One such software is 'sentence-transformers/all-mpnet-base-v2' developed by Microsoft, and the other is 'text-embedding-3-large' developed by OpenAI. To date, they represent one of the best-performing tools available (Harris et al., 2024).[17]

Both of these embedding software tools are unique in their ability to understand not only word-to-word similarity but also sentence-to-sentence similarity. If two sentences have completely different meanings, even if they use similar words, sentence embedding models will recognize them as different. In contrast, word embedding models will perceive the sentences as similar (Ul Haq et al., 2024; Zhang et al., 2024; Mandelbaum and Shalev, 2016; Li et al., 2015).

Baer and Purves (2023) demonstrates that the 'sentence-transformers/all-mpnet-base-v2' approach significantly outperforms TF-IDF in identifying similar documents, as judged by human annotators. Existing studies have predominantly relied on word

[15]SOC 17-3021: Aerospace Engineering and Operations Technologists and Technicians.

[16]SOC 25-1065: Political Science Teachers, Postsecondary.

[17]While both OpenAI's 'text-embedding-3-large' and Microsoft's 'sentence-transformers/all-mpnet-base-v2' are among the best-performing tools available, they are not the only top performers. Other models like NVIDIA's 'NV-Embed' and Salesforce's 'SFR-Embedding' also demonstrate exceptional performance (Lee et al., 2024; Meng et al., 2024).

embeddings. For instance, studies have utilized TF-IDF (Autor et al., 2024; Kogan et al., 2021; Webb, 2019) and BERT (Frugoli and ESCO, 2022). To the best of our knowledge, we are the first to apply sentence embedding technology in the field of economics.

By integrating the similarity results from both the Microsoft and OpenAI embeddings, we select the top 0.106 percent of similarity scores from the entire pool of matched pairs across years and sectors, ensuring that the matches are highly relevant. We exclude any scores below this threshold. Although this cutoff is higher than those used in other studies (Autor et al., 2024; Kogan et al., 2021; Webb, 2019), we believe it is essential to ensure the selection of highly relevant matches. Detailed reasoning and processes are provided in Appendix D.

Table 1 presents an example of these matches. Aerospace Technicians in the year 2018 are matched to multiple patents, as illustrated by the three patents included in the table. The first two patents have similarity scores above 0.660, thus meeting this cutoff criterion. However, the last patent, with a similarity score of 0.650, does not meet the threshold and is therefore excluded.

More technical descriptions of the patent weighting process are available in this link to enable accurate replication of the entire process by readers. This documentation will also help in deciding which tools to use between OpenAI and Microsoft, as each has its own advantages and disadvantages.

Recently, Autor et al. (2024) published a paper in which they claim to be the pioneers in empirically measuring innovation in human tasks. However, their weighting methodology is fundamentally different from ours. They identify newly emerged micro-occupations from the Census Alphabetical Index of Occupations and Industries (CAI), which is updated about once every decade. These new micro-occupations are weighted based on patent matching similarity scores. Their rationale is that existing patents prior to the advent of a new occupation may have catalyzed its emergence. Consequently, these patents influence the weighting assigned to the new occupation. For example, if the occupation of a fingernail artist emerged but was not preceded by any related patents, it would not be considered a labor-augmenting innovation (IHT in our terminology). Their focus is on *the number of patents* prior to the emergence of new occupations. In contrast, our emphasis on IHT involves *the annual 'Task score'*, which is adjusted by applying patent weights to the annual changes in tasks.

4.2 Innovation in Robots

The International Federation of Robotics (IFR) provides data on the number of automated robots (both flow and stock) at the country-industry-year level. Instead of using the raw data on the number of robots from the IFR, Acemoglu and Restrepo (2020) proposed utilizing the Adjusted Penetration of Robots (APR) to proxy automation. For a detailed explanation of APR, please refer to Appendix E. We will enhance this metric

Table 1: Sample Matches

SOC Descriptions for SOC 17-3021	Patent Number	Patent Descriptions	Similarity Score
Aerospace Engineering and Operations Technologists and Technicians. Operate, install, adjust, and maintain integrated computer, communications systems, consoles, simulators, and other data acquisition, test, and measurement instruments and equipment, which are used to launch, track, position, and evaluate air and space vehicles. May record and interpret test data.	9980298	This invention related to airplanes and more particularly to the use of flight deck multifunction displays to distribute various kinds of information to the flight crew of an airplane. Modern commercial airplanes include numerous avionics display systems and electronic control systems. The use of such systems is regulated and approved by various governmental authorities around the world. Such systems are classified by these regulatory authorities according to the hazard level presented to an aircraft in flight if a system fails.	0.742
	10127257	The invention generally relates to the field for monitoring the operating state of an aircraft. The invention relates to a method for creating a database of operating states of aircraft, a method for formulating a map of said operating states, and a method for monitoring the operation of an aircraft from such a map. In the field of aeronautics, it is important to be able to monitor the operating condition of an aircraft in order to predict and plan maintenance operations on the latter.	0.713
	09982534	The present disclosure relates generally to oilfield drilling and production, and more particularly, but not by way of limitation, to systems and methods for communicating information from a downhole location to a surface location. Drilling and production operations are improved with greater quantities of information relating to the conditions and drilling parameters downhole. The information is, at times, obtained by removing the drilling assembly and inserting a wireline logging tool. With great frequency today, information is obtained while drilling with measurement while drilling (MWD) or logging while drilling (LWD) techniques.	0.650

further, referring to it as the Innovation in Robots (IRB), which will serve as a proxy for dI in Equation (9) in the Regression section.

One issue with APR is that it effectively represents $d(I - N + 1)$, not dI, which is the true measure of automation. Our introduction of the proxy for dN, the IHT, as explained in the previous section, enables us to address this issue in the following manner.

From Equation (9) in the Regression section,

$d \ln S_L^f = \alpha_1 dI + \alpha_2 dN + \alpha_3 d \ln \mathbb{W} + \alpha_4 d \ln \psi + \alpha_5 d \ln R + \alpha_6 d \ln \gamma$

Therefore, on the right-hand side,

$$\begin{aligned} \alpha_1 dI + \alpha_2 dN &= \alpha_1 d(I - N + 1) + \alpha_1 dN + \alpha_2 dN \\ &= \alpha_1 \text{APR} + \alpha_1 \text{IHT} + \alpha_2 \text{IHT} \\ &= \alpha_1 (\text{APR} + \text{IHT}) + \alpha_2 \text{IHT} \\ &= \alpha_1 \text{IRB} + \alpha_2 \text{IHT} \end{aligned} \tag{7}$$

In short, we use the Innovation in Robots (IRB) to proxy automation, dI. IRB is essentially a summation of APR and IHT. One might wonder why we don't simply use dI from the beginning instead of using $d(I - N + 1) + dN$. The issue here is that there is no effective alternative to proxy dI. As mentioned earlier, the number of robots used is the result of economic equilibrium and is not the abstract concept of dI. Should readers be curious about the outcomes if the regression had employed APR instead of IRB, these results are provided in the Robustness Check section.

4.2.1 Variance Adjustment

Since APR and IHT are constructed variables, they are not directly comparable. Given that IRB is constructed by summing APR and IHT (as shown in Equation (7)), ensuring comparability between APR and IHT, especially in terms of variance, is important.

In Equation (8), the right-hand side represents the newly adjusted IHT, whereas the left-hand side details the adjustment process. Since the variances of IHT and APR are not directly comparable, we adjust by multiplying by $\frac{\sigma_{\text{APR}}}{\sigma_{\text{IHT}}}$ to equate the variance of IHT with that of APR. Here, σ represents the standard error.

$$\text{IHT} \times \frac{\sigma_{\text{APR}}}{\sigma_{\text{IHT}}} \times \frac{\sigma_{\text{inferred N}}}{\sigma_{\text{inferred (I-N)}}} \quad \Rightarrow \quad \text{IHT} \tag{8}$$

We then multiply by $\frac{\sigma_{\text{inferred N}}}{\sigma_{\text{inferred (I-N)}}}$. Here, 'inferred N' (the inferred value of IHT) and 'inferred I' (the inferred value of APR) are obtained by replicating the methodology of Acemoglu and Restrepo (2019), as detailed in Section 4.1 and Appendix A . We have extended this replication to ten countries and continued it through 2019.

While the variances of APR and IHT were not directly comparable, those of 'inferred N' and 'inferred I' are. This comparability stems from the fact that Acemoglu and Restrepo (2019) inferred these values using the same set of variables, particularly focusing on the labor share. Our approach involves adjusting the variance of IHT so that the difference in variance between IHT and APR matches that between 'inferred N' and 'inferred I'.

According to our replication, the ratio $\frac{\sigma_{\text{inferred N}}}{\sigma_{\text{inferred (I-N)}}}$ equals $\frac{54.927}{17.923} = 3.435$. Throughout this paper, we will employ the variance-adjusted version of IHT. To ensure robustness, we additionally provide regression tables in Section 7, using a ratio of $\frac{\sigma_{\text{inferred N}}}{\sigma_{\text{inferred (I-N)}}} = 1$ (i.e., with no adjustment). All other analyses remain unchanged even when using this value instead of 3.435.

4.3 Robot Price

Unfortunately, the International Federation of Robotics (IFR) no longer provides information on the prices of robots. IFR provided robot prices in the form of an average unit price until 2009, and as a price index until 2005. Klump et al. (2021) and Jurkat et al. (2022) provide in-depth information on this topic.[18] An alternative method to obtain robot prices is by following the approach of Fernandez-Macias et al. (2021), which involves the use of UN Comtrade data.[19] We adopted this method, which illustrate in their Figures 3 and A1 that the robot price trends based on IFR and UN Comtrade data are similar. Furthermore, they demonstrate that the robot price has been steadily declining.[20]

4.4 Non-robot Capital Price

Denote total capital that includes robot and non-robot as K. Also, denote robot capital and non-robot capital as M and R, respectively. Then it follows that

$$\text{gr\_Price}_K = \text{gr\_Price}_M \frac{\text{Cost}_M}{\text{Cost}_K} + \text{gr\_Price}_R \frac{\text{Cost}_R}{\text{Cost}_K}$$

, where 'gr' denotes the growth rate. The implication of this equation is that the level and scale of the prices do not matter in this growth rate relationship. The above

[18] They noted, "Due to the considerable effort involved and owing to compliance issues, the IFR no longer continues to construct the price indices."

[19] https://comtradeplus.un.org/

[20] The data generation process is as follows: UN Comtrade provides annual import and export values in dollar for 'Machinaery and mechanical appliances; industrial robot, n.e.c. or included. (HS847950)' They also provide the quantity of these values for both imports and exports. Hence, we infer the robot prices by dividing the dollar values by their quantities.

equation can be rearranged to

$$\text{gr\_Price}_R = \frac{\text{gr\_Price}_K - \text{gr\_Price}_M \times \alpha}{1 - \alpha}$$

, where α is $\frac{\text{Cost}_M}{\text{Cost}_K}$. This completes the derivation of the growth rate of price for the non-robot capital.

For the capital price, gr_Price_K, we strictly adhere to the approach outlined by Karabarbounis and Neiman (2014) throughout this paper. For detailed explanations, please refer to Appendix F. We have values for Cost_K from KLEMS data. For further explanations regarding this, please refer to Appendix G.

We can estimate Cost_M by sector and country through two approaches. The first approach employs the value obtained using the approach introduced in Section 5.3. This approach yields the ratio $\frac{\text{Robot Cost}}{\text{Labor Cost}} = 2.813\%$, and labor cost information is available from the KLEMS dataset. Consequently, we can calculate Cost_M based on this information. However, this approach is contingent on labor cost values, raising concerns that the ratio $\frac{\text{Robot Cost}}{\text{Labor Cost}} = 2.813\%$ may vary significantly across sectors and countries. Therefore, we propose an alternative approach.

The alternative approach leverages information from the alternative method detailed in Appendix H.1. In this method, we have determined the cost ratio between OMach and robots to be $13.595 : 2.149$, where 'OMach' refers to the machinery and equipment in the KLEMS. Given that we possess detailed OMach cost data by sector and country, we can subsequently estimate Cost_M. This approach circumvents the need for labor cost data. By using this approach, we complete our derivation of the growth rate of non-robot capital price, which will be used in our regression analysis.

4.5 Labor Price

For the wage variable, we utilize data from KLEMS, as detailed in Appendix G.

5 Regressions

5.1 Regression Equations

By taking the natural log of Equation (6) and then computing the total derivative of the resulting equation with respect to the exogenous variables in the model (I, N, $\mathbb{W}$, and ψ, R, and γ), we obtain the following estimating equation:

$$
\begin{aligned}
d\ln S_L^f = {} & \\
& \underbrace{\left[\underbrace{-\frac{\left(\frac{W_I}{\gamma_I}\right)^{1-\zeta}}{\int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj}}_{Ⓐ} + \underbrace{\left(-(1-\zeta) + S_K^f(1-\sigma)\right)}_{Ⓑ} \underbrace{\frac{1}{1-\zeta} \frac{\psi^{1-\zeta} - \left(\frac{W_I}{\gamma_I}\right)^{1-\zeta}}{P_T^{1-\zeta}}}_{Ⓒ} \right]}_{\alpha_1} dI \\
& + \underbrace{\left[\underbrace{\frac{\left(\frac{W_N}{\gamma_N}\right)^{1-\zeta}}{\int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj}}_{Ⓓ} + \underbrace{\left(-(1-\zeta) + S_K^f(1-\sigma)\right)}_{Ⓑ} \underbrace{\frac{1}{1-\zeta} \frac{-\psi^{1-\zeta} + \left(\frac{W_N}{\gamma_N}\right)^{1-\zeta}}{P_T^{1-\zeta}}}_{Ⓔ} \right]}_{\alpha_2} dN \\
& + \underbrace{\left[(1-\zeta) + \left(-(1-\zeta) + S_K^f(1-\sigma)\right) S_L^T \right]}_{\alpha_3} d\ln \mathbb{W} \\
& + \underbrace{\left[\left(-(1-\zeta) + S_K^f(1-\sigma)\right) S_M^T \right]}_{\alpha_4} d\ln \psi \\
& \underbrace{- \left[S_K^f(1-\sigma) \right]}_{\alpha_5} d\ln R. \\
& \underbrace{- \left[(1-\zeta) + \left(-(1-\zeta) + S_K^f(1-\sigma)\right) S_L^T \right]}_{\alpha_6} d\ln \gamma
\end{aligned}
\tag{9}
$$

, where S_L^f represents labor share times markup, I is automation, N is innovation in human tasks, ψ is robot price, R is non-robot capital price, and γ is labor productivity. $\mathbb{W} \equiv \frac{\int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj}{\int_I^N W_j^{-\zeta} \gamma_j^{\zeta-1} dj}$ is the average wage, and assume $d\ln\mathbb{W} = d\ln W_j$ for all j. Additionally, $d\ln\gamma$ represents the change in labor productivity. It is also assumed that $d\ln\gamma = d\ln\gamma_j$ for all j.

S_M^T (S_L^T) represents the share of robot cost (labor cost) in the total combined task cost, which comprises both labor and robot costs. By definition, $S_M^T + S_L^T$ equals one. In detail, these are described mathematically as follows:

$$S_M^T = \frac{(I - N + 1)\psi^{1-\zeta}}{P_T^{1-\zeta}}$$

$$S_L^T = \frac{\int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj}{P_T^{1-\zeta}}$$

$$\text{, where } P_T^{1-\zeta} = (I - N + 1)\psi^{1-\zeta} + \int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj.$$

The coefficients of the six explanatory variables ($d\ln\gamma$, dI, dN, $d\ln\mathbb{W}$, $d\ln R$, and $d\ln\psi$) in Equation (9) reflects not only the direct effect caused by the change in the variable, but also the general equilibrium effects that influence the labor share through changes in the price of the aggregated tasks.

Based on the specification in Equation (9), we provide consistent regression equations as below:

$$\begin{aligned} \text{gr\_(laborshare} \times \text{markup)} =& \alpha_1 \text{IRB} + \alpha_2 \text{IHT} \\ &+ \alpha_3 \text{gr\_labor price} + \alpha_4 \text{gr\_robot price} \\ &+ \alpha_5 \text{gr\_non-robot capital price} \\ &+ \alpha_6 \text{gr\_labor productivity} \\ &+ \lambda_i + \lambda_j + \lambda_t + \lambda_{ij} + \varepsilon_{ijt}. \end{aligned} \tag{10}$$

gr indicates the variables are in a 5-year growth rate, and i, j, and t correspond to country, industry, and year, respectively. IRB and IHT stand for the growth rate of Innovation in Robots and Innovation in Human Tasks, respectively. We exclude the notation of *gr* from IRB and IHT, as by definition, they already represent a 5-year growth rate.

5.2 Regression Results

As outlined in Section 4, our primary specification utilizes weights based on US patents, which are embedded using softwares developed by Microsoft and OpenAI. To assess robustness, we also present regression tables and accounting figures derived from the unweighted version, and the version weighted by using wages, as detailed in the Robustness Check section.

We present our Ordinary Least Squares (OLS) results in Table 2. Standard errors are clustered by country and sector to account for the serial correlation. To improve readability, both the coefficients and standard errors have been multiplied by 100. Columns (1) through (3) include labor productivity as an explanatory variable in the regressions, while Columns (4) and (5) do not. In this study, labor productivity is defined as value-added per worker ($\frac{Y}{L}$), which may be a rough measurement. To ensure validity, we present regression results excluding this variable in Columns (4) and (5). In these columns, we interpret that the fixed effects capture the labor productivity.

Meanwhile, the sum of the coefficients for $d\ln\mathbb{W}$, $d\ln\psi$, and $d\ln R$ equals zero (i.e., $\alpha_3+\alpha_4+\alpha_5=0$). Similarly, the sum of the coefficients for $d\ln\mathbb{W}$ and $d\ln\gamma$ also equals zero (i.e., $\alpha_3+\alpha_6=0$). Accordingly, we apply these restrictions in our analysis as follows: Columns (1) and (4) are presented without the coefficient restrictions; Columns (2) and (5) display the OLS results with only the first restriction ($\alpha_3+\alpha_4+\alpha_5=0$); Column (3) displays the OLS results incorporating both restrictions ($\alpha_3+\alpha_4+\alpha_5=0$ and $\alpha_3+\alpha_6=0$).

Table 2: Regressions

	With labor productivity			Without labor productivity	
	(1)	(2)	(3)	(4)	(5)
Restrction 1	No	Yes	Yes	No	Yes
Restrction 2	No	No	Yes	No	NA
α_1: IRB	-0.089***	-0.093***	-0.118***	-0.071***	-0.056**
	(0.022)	(0.024)	(0.020)	(0.020)	(0.024)
α_2: IHT	0.108***	0.111***	0.143***	0.090***	0.075***
	(0.024)	(0.026)	(0.021)	(0.022)	(0.026)
α_3: gr_labor price	14.801***	16.321***	4.574***	11.873***	11.595***
	(3.512)	(3.512)	(1.599)	(2.996)	(3.367)
α_4: gr_robot price	0.007	-0.042	0.732	0.162	0.243
	(0.940)	(0.944)	(1.008)	(0.953)	(0.983)
α_5: gr_non robot capital price	-19.205***	-16.280***	-5.306***	-18.668***	-11.838***
	(3.353)	(3.267)	(1.679)	(3.360)	(3.161)
α_6: gr_labor productivity	-2.982**	-4.665***	-4.574***		
	(1.395)	(1.397)	(1.599)		
N	998	998	998	998	998
R^2	0.655			0.650	

The coefficients and the standard errors have been multiplied by 100 for better readability.
Standard errors in parenthesis are clustered by country and sector.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

In assessing the congruence between the regression results and the model's predictions, two findings are noteworthy. First, the model delineates the coefficient for robot price as α_4, with the term $S_M^T = 2.81\%$ included, which we estimated in Section 5.3. The model thus anticipates this coefficient to be of an insignificantly small value. In line with this prediction, the regression coefficient for robot price is not statistically significant, and the point estimate lacks precision. Second, the OLS results maintain consistency in both magnitude and direction, regardless of whether the restriction is applied. Utilizing OLS without the restriction (as shown in Column (1)), we test the null hypothesis that the restrictions are non-binding. The hypotheses for Restriction 1 and Restriction 2 are both rejected at the 0.05 significance level. This suggests a certain degree of misalignment between the data and the model's predictions. In subsequent analyses, we refer to the OLS results from Column (1) as our baseline.

5.3 Estimation of S_M^T

S_M^T represents the share of robot cost in the total combined task cost, which comprises both labor and robot costs. This metric is vital for our analysis in the Regression section. Unfortunately, no official data is available that directly quantifies this value, requiring us to rely on multiple sources for an accurate estimation.

For a detailed explanation of how we estimated S_M^T, please refer to Appendix H. By synthesizing all available information, we estimate S_M^T to be 2.813% for the total manufacturing sectors. An alternative method detailed in Appendix H.1 estimates the S_M^T value at 2.104%. However, we consider the method outlined in this section to be more accurate and reliable, leading us to conclude that the S_M^T value is 2.813%.

5.4 Estimation of σ and ζ

By utilizing Equation (9) along with the regression results, we estimate the values of σ and ζ. σ represents the elasticity of substitution between the aggregate task and non-robot capital. Notably, labor costs account for 97.2% of the aggregate task cost, while non-robot capital accounts for 91.1% of the 'overall' capital cost. Thus, σ serves as a close proxy for the elasticity of substitution between labor and overall capital.

The literature on the elasticity of substitution between labor and 'overall' capital is extensive. However, relatively less attention has been given to ζ, which is the elasticity of substitution between tasks in the model but can also be interpreted as the elasticity of substitution between human workers and robots at an aggregate level.[21] Furthermore, to our knowledge, no previous studies have attempted to estimate both the elasticity of substitution between labor and non-robot capital and between labor

[21] In Acemoglu and Restrepo (2019), the model does not distinguish between robot and non-robot capital, using only overall capital. Consequently, their measure of the elasticity of substitution between tasks is interpreted as the elasticity between human workers and overall capital at an aggregate level.

and robot capital within a single framework. One contribution of this section is to provide such estimates.

We detail the methodology for estimating these two elasticities, σ and ζ in Appendix I. Our results are as follows: first, we calculate $\sigma = 0.611$, with a 95% confidence interval for σ of $(0.477, 0.746)$. σ differs from the elasticity of substitution between labor and non robot-capital, but as mentioned, σ serves as a close proxy of this elasticity. In Appendix K, we provide a formal estimation of the elasticity of substitution between labor and non-robot capital using the estimation of σ. This measure closely aligns with the measures used by Karabarbounis and Neiman (2014) and Glover and Short (2020), and our estimate ranges between 0.611 and 0.722. Thus, this result contributes to literature by providing additional empirical evidence that the elasticity of substitution between labor and non-robot capital is less than one, indicating a gross complementary relationship between the two. This is supported by most literature, as suggested by Chirinko (2008), Grossman and Oberfield (2022), and Glover and Short (2020).

We also estimate $\zeta = 2.374$, with a 95% confidence interval ranging from 1.017 to 3.730. This result is similar to, but slightly higher than, the findings of DeCanio (2016), which suggest a ζ of about 1.9. Importantly, this estimate indicates that ζ is greater than one, highlighting how improvements in robot productivity —reflected by a decrease in robot price— affect the labor share. We discuss this in further detail in Section 5.6.

Additionally, we perform a Wald test on $\zeta - \sigma$ and obtain a 95% confidence interval ranging from 0.378 to 3.147. This indicates that ζ is statistically different from and greater than σ, reinforcing the importance of distinguishing between robot and non-robot capital when analyzing the impact of robots on the labor share.

5.5 Effects of Automation and Innovation in Human Tasks on Labor Share

The regression results show that advance in automation negatively affects labor share and innovation in human tasks positively affects labor share. These results are consistent with many other studies such as Acemoglu and Restrepo (2022a) and Autor et al. (2024). When interpreting these results, the direct effects of these two innovations are theoretically straightforward: automation directly reduces the labor share as robots replace human labor, and innovations in human tasks increase the labor share as labor takes on roles previously performed by robots. These direct effects are represented by Ⓐ and Ⓓ in Euqation (9), respectively.

Interestingly, this paper also elucidates how these innovations indirectly affect the labor share, contributing to the existing literature. The indirect effects are defined as follows: First, automation and innovations in human tasks alter the composition of tasks performed by robots and humans. Second, this change in task composition affects the relative price of labor compared to robots and the aggregate task price. Finally, these changes influence the labor share through substitution among labor, robots, and non-robot capital. These indirect effects are captured by the second term in the coefficients of dI and dN, which are Ⓑ $\times$ Ⓒ and Ⓑ $\times$ Ⓔ, respectively.

To analyze the indirect effects of these two innovations, we first provide an intuitive explanation for the analytic expressions (Ⓑ$\times$Ⓒ and Ⓑ$\times$Ⓔ). Ⓒ and Ⓔ represent the relative reduction in the price of aggregated tasks compared to non-robot capital due to innovation. For instance, during automation, robots assume task I from labor, thereby reducing the production cost of aggregated tasks relative to non-robot capital, as expressed by $\frac{1}{1-\zeta}\frac{-\psi^{1-\zeta}+\left(\frac{W_N}{\gamma_N}\right)^{1-\zeta}}{P_T^{1-\zeta}}$. When the price of aggregated tasks changes by 1, while other determinants remain constant, the labor share changes by $-(1-\zeta)$ due to the substitution between labor and robots within aggregated tasks, and by $S_K^f(1-\sigma)$ due to the substitution between aggregated tasks and non-robot capital. Overall, the effect of the aggregated task price is determined by the multiplication of the change in the price of the aggregated task, which is Ⓒ for automation and Ⓔ for innovation in human tasks, and the sum of $-(1-\zeta)$ and $S_K^f(1-\sigma)$, represented by Ⓑ.

Therefore, the term $-(1-\zeta)+S_K^f(1-\sigma)$, which is determined by the values of ζ and σ, plays a crucial role in the indirect effects. To elucidate the sign of this term, we substitute the estimates for σ and ζ obtained from the regression results in Column (1) and find that this term evaluates to $1.566 > 0$. Furthermore, to test its significance, we employ stochastic variables for σ and ζ and conduct a Wald test on the null hypothesis that $-(1-\zeta)+S_K^f(1-\sigma)=0$, yielding a confidence interval of $(0.197, 2.934)$ at the 0.05 significance level. This test allows us to reject the null hypothesis, supporting a reasonable inference that the term $-(1-\zeta)+S_K^f(1-\sigma)$ is indeed positive.

Lastly, to determine how indirect effects operate for both automation and innovation in human tasks, we analyze the signs of Ⓒ and Ⓔ, respectively. We provide a detailed proof in Appendix J for this analysis. Given that Ⓒ is less than zero and Ⓔ is greater than zero, we conclude that the negative effects of automation and the positive effects of innovation in human tasks are both amplified through their indirect effects.

5.6 Effects of Price Factors on Labor Share

5.6.1 Labor Price and Non-Robot Capital Price

The regression findings highlight the impact of factor prices on labor share. They reveal a positive relationship between the price of labor and labor share. Based on α_3, this result can be interpreted as follows: The robot cost share, denoted by S_M^T, is a very small value, specifically 0.028. This indicates that when wages change, substitution between labor and robots does not have a significant effect, and substitution between labor and non-robot capital plays a more important role, as demonstrated below:

$$\begin{aligned}
\alpha_3 =& (1-\zeta) + \Big(-(1-\zeta) + S_K^f(1-\sigma)\Big) S_L^T \\
=& (1-\zeta)(1-S_L^T) + S_K^f(1-\sigma)S_L^T \\
=& -0.039 + S_K^f(1-\sigma)S_L^T \\
\approx&\ S_K^f(1-\sigma)S_L^T = 0.14801 > 0.
\end{aligned}$$

Similarly, the regression suggests a negative association between the price of non-robot capital and labor share, which can be explained with the equation below.

$$\alpha_5 = -\Big[S_K^f(1-\sigma)\Big] < 0 \tag{11}$$

The two aforementioned results stem from the gross complementarity between labor and non-robot capital. Specifically, when wages rise, employment levels do not decrease proportionally, leading to an increase in labor share. Similarly, an increase in the price of non-robot capital results in a decline in labor share.

In Figure 10, provided in Appendix M, we replicate the derivation of capital price following the approach used by Karabarbounis and Neiman (2014) (hereafter referred to as KN), utilizing the KLEMS data version. This ensures that the 'overall' capital price variable is identical to that used by KN. Subsequently, we derive the non-robot capital price variable as detailed in Section 4.4. This non-robot capital price variable is then consistently utilized throughout Sections 5 and 6. Our data indicate that the prices of non-robot capital have generally increased over the past 15 years, as illustrated in Figure 10 in Appendix M. This observation might initially appear contradictory to the claims of KN, who reported a rapid global decline in capital prices (see Figure 7 of their paper). However, our Figure 10 is consistent with their findings, considering that capital prices began to rise from around year 2000. Furthermore, their figure aggregates data from all countries worldwide, whereas our analysis is more focused, presenting data at the country level for only ten selected countries. Given the negative coefficient of non-robot capital price, this indicates a decline in labor share.

5.6.2 Robot Price

The regression results indicate a positive, albeit small, association between robot price and labor share. This insignificance is attributed to the low share of the robot cost ($S_M^T = 2.8\%$). This means that even if robot prices change, their impact on labor share will inevitably be small.

$$\alpha_4 = \Big(-(1-\zeta) + S_K^f(1-\sigma)\Big) S_M^T > 0 \tag{12}$$

The positive correlation is primarily dependent on the condition $-(1-\zeta)+S_K^f(1-\sigma)>0$, which is fundamentally attributed to $\zeta>1$ and $\sigma<1$. The underlying logic is as follows: (1) Since robots and labor are gross substitutes ($\zeta>1$), a decrease in robot prices directly reduces the labor share. (2) A decrease in robot prices also causes a reduction in the price of aggregated tasks. Given that aggregated tasks and non-robot capital are gross complements ($\sigma<1$), this leads to a further decrease in labor share. These results hold significant implications as they empirically substantiate the theoretical conditions identified in other literature as necessary for scenarios where improvements in robot productivity lead to a decrease in labor share. For example, Berg et al. (2018), which distinguishes between robot and non-robot capital and employs the same two-tier CES structure as our model, stipulates a condition similar to $\alpha_4>0$ for a decrease in labor share to occur not only in the long term but also in the short term when robot productivity increases. It is important to note that, to satisfy either $\alpha_4>0$[22] in our model or the condition mentioned in Berg et al. (2018), it is necessary that the elasticity of substitution between labor and robot significantly exceeds the elasticity of substitution between aggregated tasks and non-robot capital.

The positive relationship with robot prices in our model uncovers two pivotal mechanisms that impact labor share as advancements in robotics occur. First, enhanced robotic capabilities allow for the execution of tasks previously exclusive to humans, thereby reducing labor share. Second, a decline in the price of robots, without a corresponding enhancement in functionality, also exerts a negative impact on the labor share.

In the future, we anticipate that the coefficient for robot price will become more prominent, yielding a positive association as the share of robots in society increases. This expectation is attributable to the term S_M^T, the share of robot costs. It is important to note that, among the three price factors in Equation (9), S_M^T is uniquely associated with the price of robots.

5.7 Steady State Analysis

In this subsection, we present a regression table based on the premise that removing short-run fluctuations may lessen the endogenous responses they exhibit. This approach aligns with the one taken by Karabarbounis and Neiman (2014), who removed the year dimension from variables in their main analysis, as demonstrated in their Equation (19). For each variable, they calculated a coefficient β from the fitted line as follows:

$$Y_{ij}=\alpha+\beta T_{ij}$$

[22]We can rearrange this inequality as follows: $\zeta>1-S_K^f+S_K^f\sigma$

, where ij represents different groups, and T_{ij} denotes the year, which has continuous values. The coefficient β represents the linear trend within each group ij, serving as a growth rate in the steady state. Consequently, this approach significantly reduces the number of observations. Column (1) of Table 3 is the regression result using the baseline model (i.e., Column (1) of Table 2). With only 97 observations, the signs of the variables remain unchanged compared to the baseline results in Table 2.

To increase the number of observations, we employ the Hodrick-Prescott filter, setting the smoothing parameter to 10, which is higher (thus smoother) than the default value of 6.25 for yearly data. This approach prevents the data from being completely flattened to a line, allowing us to maintain the same number of observations as in the baseline regression data. An important aspect of this analysis is the use of clustered standard errors, as serial correlation is highly anticipated following the smoothing process. Columns (2) of Table 3 is the result using the baseline model.

Table 3: Steady State Regressions

	KN Line Trend	HP-filter
	(1)	(2)
IRB	-0.068**	-0.107***
	(0.027)	(0.036)
IHT	0.212***	0.108***
	(0.027)	(0.039)
gr_labor price	16.992*	12.155**
	(9.189)	(5.165)
gr_robot price	16.413	-1.123
	(9.725)	(1.048)
gr_non robot capital price	-7.565	-22.749***
	(8.932)	(5.289)
gr_labor productivity	-18.001**	0.226
	(5.975)	(3.002)
N	97	997
R^2	0.836	0.669

The coefficients and the standard errors have been multiplied by 100 for better readability.
Column (1): Standard errors in parenthesis are clustered by sector.
Column (2): Standard errors in parenthesis are clustered by country and sector.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The significance of the flattening approach discussed in this section lies in its assumption that high-frequency fluctuations signify an endogenous response. By identifying a steady state trend, this method can mitigate the issue of endogeneity, albeit to a limited extent. To explore the implications of the regression results further, we will now shift to the accounting exercise.

6 Accounting Exercise

Based on the main regression results from Column (1) in Table 2, we have generated Figures 3, 4, and 5. In this paper, we exclusively focus on country-level variation to maintain brevity. Accordingly, the values in these figures are derived by aggregating data at the country level. During this aggregation process, 'Average variables' are consolidated by weighting the value-added in each sector and year. Intuitively, the values illustrated in Figures 3, 4, and 5 quantify the extent to which each factor influences the growth rate of the markup-adjusted labor share (S_L^f).[23]

Figure 3: Labor shares

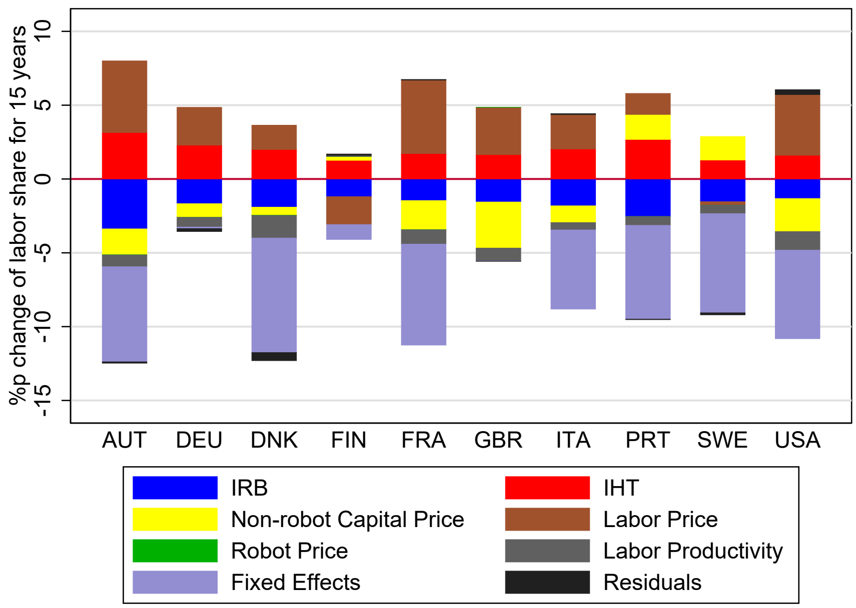


Figure 3 illustrates accounting results, showing that IRB and IHT are important factors affecting the labor share. Specifically, an increase in IRB leads to a decrease in labor share, while an increase in IHT leads to an increase. This confirms Autor (2015)'s argument that "the sustained relevance of human labor in the future will largely depend on the pace at which 'innovation in human tasks' outstrips the advancement of automation."

Our data indicate that the prices of non-robotic capital have generally risen over the past 15 years, as illustrated in Figure 10 provided in Appendix M. The negative coefficient associated with the price of non-robotic capital suggests a corresponding

[23] S_L^f is defined in Equation (6) in the Model section.

Figure 4: Labor shares (NET of IRB and IHT)

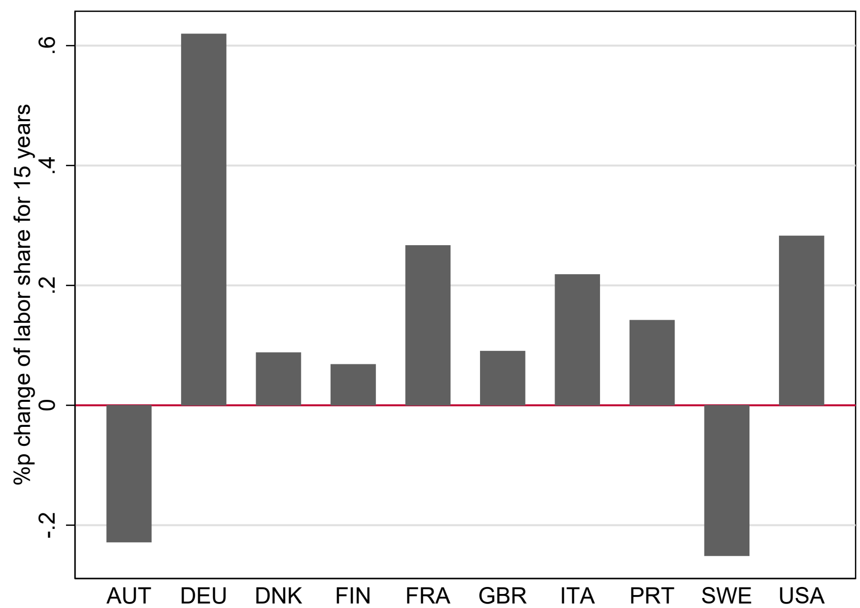
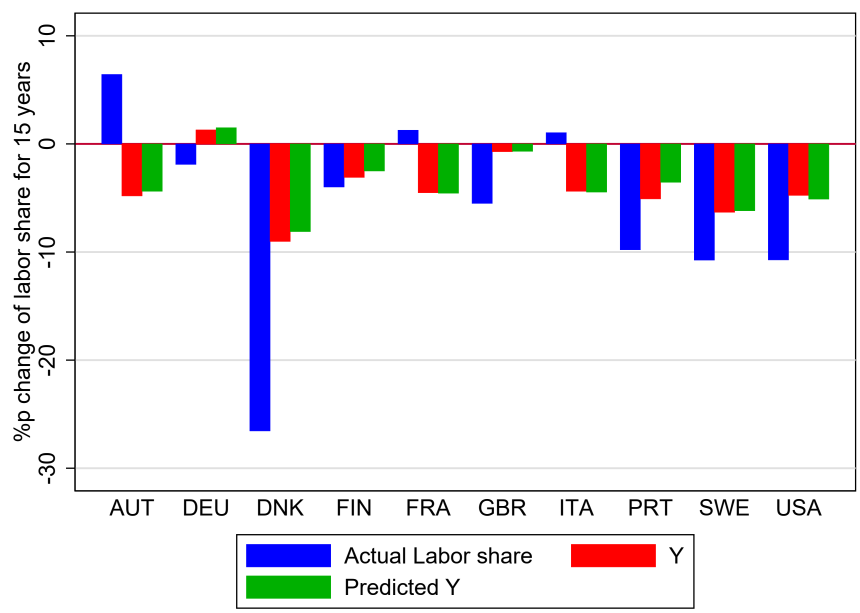


Figure 5: Labor shares (Zoom out)



decline in the labor share. This observation stands in contrast to the argument presented by Karabarbounis and Neiman (2014). Given that the coefficients for both are positive, this indicates that labor prices have increased globally during the period of our study, while robot prices have decreased. Finally, Figure 3 presents some portion of fixed effects.

Figure 4 differs from Figure 3 in that it nets out IRB and IHT, thereby canceling out effects in opposite directions. Despite their largely offsetting effects, the overall result is that innovation in human tasks surpasses the impact of automation.

Figure 5 shows even more zoomed-out view. Here, Y represents the left-hand side of the regression equation, namely, the markup-adjusted labor share (S_L^f). The term 'Predicted Y' refers to Y minus the residuals. The 'Actual Labor Share' is calculated without the markup adjustment. Notably, there is a significant difference between the 'Actual Labor Share' and Y. This discrepancy is primarily attributed to the markup. For instance, in the USA, this gap is substantially negative, indicating that the markup has had a negative impact on the labor share

The negative impact of markup on labor share is comprehensively detailed by Autor et al. (2020). They demonstrated that an increase in market concentration in the USA has led to a rise in markup. Given the negative correlation between markup and labor share (a coefficient of -1), the increased concentration has consequently resulted in a decrease in labor share.

7 Robustness Check

To conserve space, we have relocated the tables and figures relevant to this section to Appendix M. In all tables in the Robustness Check section, the signs, significances, and coefficient values are consistent with those in the main Regression section. Recognizing that a precise definition of IHT encompasses the development of tasks not inherently limited to those related to patents, we include an unweighted version in Table 5. Across all columns, the signs, significances, and coefficient values remain consistent with those in the main table in the Regression section, supporting the conclusions drawn in Section 5 even without patent weighting.

As an alternative to patent weighting, we can use wages as weights. For instance, in 2010, a task performed by a nuclear scientist, who earns three times the wage of a fingernail artist (just an example), is valued three times more than a task performed by a fingernail artist. Table 6 is the result.

Section 4.2.1 describes our use of a variance-adjusted version of IHT, represented in Equation (8). Table 7 presents an unadjusted version of IHT to demonstrate that our results do not stem from selective adjustments.

Meanwhile, as outlined in Section 4.2, IRB is essentially a summation of APR and IHT. While previous tables have used IRB, we also include Table 8 with APR only, following the specification used by Acemoglu and Restrepo (2020) for automation.

This table employs a variance-adjusted version of IHT and is patent weighted. For additional variants, such as APR, unadjusted IHT, and wage-weighted IHT, please see the replication code.

Additionally, we provide accounting figures corresponding to the various versions discussed. Initially, our main specification in the Regression section was presented in Table 2, with its corresponding accounting figure shown in Figure 3. Subsequently, we introduced an unweighted version, presented in Table 5, with its corresponding accounting figure depicted in Figure 11. Third, a wage-weighted version was presented in Table 6, and its corresponding accounting figure is shown in Figure 12. Fourth, we provided a variance-unadjusted version of IHT in Table 7, with the corresponding accounting figures in Figure 13. Finally, Figure 14 provides an accounting figure based on the steady steady analysis provided in Section 5.7. Based on the figures and tables presented in the Robustness section, we conclude that all our intuitions and results from the main regression section remain stable across various specifications.

Lastly in this section, Table 9 provides the point estimates and 95% confidence intervals for ζ, σ, and $-(1-\zeta)+S_K^f(1-\sigma)$ for each variation discussed above. The results are consistent with the main specifications discussed in the Regression section, allowing us to assert confidently that the implications derived from these stochastic variables are robust.

8 Concluding Remarks

In summary, this paper aims to unravel the factors contributing to the recent downtrend in labor share, placing a special emphasis on the roles of automation and innovation in human tasks. While existing literature presents a mosaic of conflicting viewpoints (Acemoglu and Restrepo, 2020; Graetz and Michaels, 2018; Dauth et al., 2021; De Vries et al., 2020; Humlum, 2019), our empirical analysis corroborates the adverse impact of automation on labor share.

Autor et al. (2024) is the first to empirically estimate 'innovation in human tasks.' We developed a distinctive approach compared to theirs to empirically measure this factor and investigate its impact on labor share. Notably, we construct cross-country data for innovation in human tasks, enabling us to investigate how this innovation affects labor share across countries. Our findings suggest that this factor effectively mitigates the negative repercussions of automation on labor share in the countries studied.

Our quantified estimates indicate that the elasticity of substitution *between labor and non-robot capital* is below one, while the elasticity of substitution *between labor and robot capital* is greater than one. These estimates facilitate a nuanced understanding of how factor prices —namely, labor, robots, and non-robot capital— affect labor share. Specifically, we observe that both the negative effect of automation and the positive effect of innovation in human tasks are amplified through the aggregated task price channel: First, automation and innovation in human tasks alter the composition of tasks performed by robots and those performed by labor. Second, this change in composition affects the aggregate task price. Finally, the change in the aggregate task price, in turn, affects labor share through substitution among labor, robots, and non-robot capital.

Our regression results suggest a positive correlation between wages and labor share, and a negative correlation between the price of non-robot capital and labor share. The underlying intuition stems from the gross complementarity between labor and non-robot capital. In this regard, this paper contributes to the literature by providing an additional empirical evidence on the gross complementarity between labor and non-robot capital.

The regression results posit a positive, albeit small, correlation between the price of robots and labor share. This implies that a decrease in robot prices is associated with a reduction in labor share. The weak nature of this correlation can be attributed to the currently minor contribution of robot costs to total costs. Moreover, based on the estimation results that the elasticity of substitution between labor and robot capital is much larger than the elasticity of substitution between labor and non-robot capital, our model anticipates that a reduction in robot prices or an improvement in robot productivity will significantly decrease labor share in the future when robots are expected to be used more in production. This highlights a different channel through which robotic technology affects labor share, apart from automation. These results are important as they empirically support the theoretical conditions in other studies that improvements in robot productivity negatively affect labor share.

Meanwhile, we would like to clarify that the focus of this paper is not to investigate whether this decline in labor share exacerbates income inequality or necessitates policy interventions. Although some studies have posited a correlation between a declining labor share and increasing income inequality, a more comprehensive examination of causality is necessary (ILO and KIEP, 2015; Torres et al., 2011). As such, we set these topics aside and concentrate on identifying the reasons for the decline within a unified framework.

However, as a policy recommendation, we suggest that governments implement ONET programs aimed at keeping people updated on task requirements for specific occupations. Providing such information will enable individuals to identify emerging labor demands and prepare accordingly, thus improving the alignment between labor supply and demand. With the advent of automated robots and artificial intelligence, there is significant discussion about the potential end of traditional labor and the emergence of a new class that may become obsolete in the workforce. Among these discussions, the emphasis on the importance of retraining in this new era is particularly noteworthy. While the USA is the only country currently offering ONET, the EU has recently initiated a similar project.[24] However, many countries, such as South Korea with its Korea Employment Information Service (KELS), offer job information and matching services but lack ONET-style service.

In the current landscape, our paper shows that while automation contributes to a declining labor share, innovation in human tasks exerts a significantly more positive impact on labor share. Drawing on our general equilibrium model, we anticipate that in the future, the robot price channel will gain greater importance as the prevalence of robot usage increases.

[24] The European Commission has recently initiated a project akin to ONET, named 'European Skills, Competences, Qualifications, and Occupations' (ESCO). ESCO has disclosed the tasks required for workers for a single year and has yet to release a Task score.

References

Acemoglu, D., C. Lelarge, and P. Restrepo (2020). Competing with robots: Firm-level evidence from France. In *AEA Papers and Proceedings*, Volume 110, pp. 383–388. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Acemoglu, D. and P. Restrepo (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American economic review 108*(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives 33*(2), 3–30.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy 128*(6), 2188–2244.

Acemoglu, D. and P. Restrepo (2022a). Demographics and automation. *The Review of Economic Studies 89*(1), 1–44.

Acemoglu, D. and P. Restrepo (2022b). Tasks, automation, and the rise in US wage inequality. *Econometrica 90*(5), 1973–2016.

Autor, D., C. Chin, A. Salomons, and B. Seegmiller (2024). New frontiers: The origins and content of new work, 1940–2018. *The Quarterly Journal of Economics*, qjae008.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics 135*(2), 645–709.

Autor, D. H. (2015, June). Why are there still so many jobs? the history and future of workplace automation. In *Journal of Economic Perspectives*, Volume 29, pp. 3–30. American Economic Association.

Baer, M. F. and R. S. Purves (2023, March). Identifying Landscape Relevant Natural Language using Actively Crowdsourced Landscape Descriptions and Sentence-Transformers. *KI - Künstliche Intelligenz 37*(1), 55–67.

Berg, A., E. F. Buffie, and L.-F. Zanna (2018). Should we fear the robot revolution?(The correct answer is yes). *Journal of Monetary Economics 97*, 117–148.

Bergholt, D., F. Furlanetto, and N. Maffei-Faccioli (2022). The decline of the labor share: New empirical evidence. *American Economic Journal: Macroeconomics 14*(3), 163–98.

Chirinko, R. S. (2008). σ: The long and short of it. *Journal of Macroeconomics 30*(2), 671–686.

Dauth, W., S. Findeisen, J. Suedekum, and N. Woessner (2021). The adjustment of labor markets to robots. *Journal of the European Economic Association 19*(6), 3104–3153.

De Vries, G. J., E. Gentile, S. Miroudot, and K. M. Wacker (2020). The rise of robots and the fall of routine jobs. *Labour Economics 66*, 101885.

DeCanio, S. J. (2016). Robots and humans–complements or substitutes? *Journal of Macroeconomics 49*, 280–291.

Fernandez-Macias, E., D. Klenert, and J.-I. Anton (2021). Not so disruptive yet? Characteristics, distribution and determinants of robots in Europe. *Structural Change and Economic Dynamics 58*, 76–89.

Frugoli, P. and ESCO (2022). The crosswalk between ESCO and O*NET (Technical Report).

Glover, A. and J. Short (2020). Can capital deepening explain the global decline in labor's share? *Review of Economic Dynamics 35*, 35–53.

Gouma, R. and M. Timmer (2013). World KLEMS Growth and Productivity Accounts Japan: Sources and notes. *Groningen Growth and Development Centre 8*.

Graetz, G. and G. Michaels (2018). Robots at work. *Review of Economics and Statistics 100*(5), 753–768.

Gregory, T., A. Salomons, and U. Zierahn (2016). Racing with or against the machine? Evidence from Europe. *Evidence from Europe (July 15, 2016). ZEW-Centre for European Economic Research Discussion Paper* (16-053).

Grossman, G. M. and E. Oberfield (2022). The elusive explanation for the declining labor share. *Annual Review of Economics 14*, 93–124.

Gutiérrez, G. and S. Piton (2020). Revisiting the global decline of the (non-housing) labor share. *American Economic Review: Insights 2*(3), 321–338.

Harris, N., A. Butani, and S. Hashmy (2024, April). Enhancing Embedding Performance through Large Language Model-based Text Enrichment and Rewriting.

Hubmer, J. and P. Restrepo (2021). Not a typical firm: The joint dynamics of firms, labor shares, and capital–labor substitution. Technical report, National Bureau of Economic Research.

Humlum, A. (2019). Robot adoption and labor market dynamics. *Princeton University*.

ILO and KIEP (2015). Inequality in G20 countries: Causes, impacts, and policy responses. *G20 Employment Working Group, Cappadocia, Turkey*.

Jorgenson, D. W. (1963). Capital theory and investment behavior. *The American Economic Review 53*(2), 247–259.

Jurkat, A., R. Klump, and F. Schneider (2022). Tracking the Rise of Robots: The IFR Database. *242*(5-6), 669–689.

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

Khan, A., Q. Shah, M. I. Uddin, F. Ullah, A. Alharbi, H. Alyami, and M. A. Gul (2020, August). Sentence Embedding Based Semantic Clustering Approach for Discussion Thread Summarization. *Complexity 2020*, 1–11.

Klump, R., A. Jurkat, and F. Schneider (2021). Tracking the rise of robots: A survey of the IFR database and its applications.

Kogan, L., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2021). *Technology-Skill Complementarity and Labor Displacement: Evidence from Linking Two Centuries of Patents with Occupations.* National Bureau of Economic Research.

Kogan, L., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2023). Technology and labor displacement: Evidence from linking patents with worker-level data. Technical report, National Bureau of Economic Research.

Lee, C., R. Roy, M. Xu, J. Raiman, M. Shoeybi, B. Catanzaro, and W. Ping (2024, May). NV-Embed: Improved Techniques for Training LLMs as Generalist Embedding Models.

Li, S., J. Zhu, and C. Miao (2015, August). A Generative Word Embedding Model and its Low Rank Positive Semidefinite Solution.

Lybbert, T. J. and N. J. Zolas (2014). Getting patents and economic data to speak to each other: An 'algorithmic links with probabilities' approach for joint analyses of patenting and economic activity. *Research Policy 43*(3), 530–542.

Mandelbaum, A. and A. Shalev (2016, October). Word Embeddings and Their Use In Sentence Classification Tasks.

Martinez, J. (2018). Automation, growth and factor shares. In *2018 Meeting Papers*, Volume 736. Society for Economic Dynamics.

Mazandu, G. K., K. Opap, F. Makinde, V. Nembaware, F. Agamah, C. Bope, E. R. Chimusa, A. Wonkam, and N. J. Mulder (2021). An Integrated Platform Supporting Semantic Similarity Score Calculation and Reproducibility.

Meng, R., Y. Liu, S. R. Joty, C. Xiong, Y. Zhou, and S. Yavuz (2024). SFR-embedding-mistral: Enhance text retrieval with transfer learning. *Salesforce AI Research Blog 3*.

Montobbio, F., J. Staccioli, M. E. Virgillito, and M. Vivarelli (2021). Labour-saving automation and occupational exposure: A text-similarity measure. Technical report, LEM Working Paper Series.

Müller, C. (2022). World Robotics 2022 - Industrial Robots.

National Center for O*NET Development (2023). O*NET Resource Center.

Oberfield, E. and D. Raval (2021). Micro data and macro technology. *Econometrica 89*(2), 703–732.

Piketty, T. and G. Zucman (2014). Capital is back: Wealth-income ratios in rich countries 1700–2010. *The Quarterly journal of economics 129*(3), 1255–1310.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2020). IPUMS USA: Version 10.0 [dataset]. *Minneapolis, Mn: Ipums 10*, D010.

Stehrer, R., A. Bykova, K. Jäger, O. Reiter, and M. Schwarzhappel (2019). Industry level growth and productivity data with special focus on intangible assets. *Vienna Institute for International Economic Studies Statistical Report 8*.

Torres, R., International Labour Organisation, and International Institute for Labour Studies (2011). World of Work Report 2011. Making markets work for jobs. *International Labour Organisation. International Institute for Labour Studies.*

Ul Haq, I., M. Pifarré, and E. Fraca (2024, January). Novelty Evaluation using Sentence Embedding Models in Open-ended Cocreative Problem-solving. *International Journal of Artificial Intelligence in Education.*

Webb, M. (2019). The impact of artificial intelligence on the labor market. *Available at SSRN 3482150*.

Zhang, B., K. Chang, and C. Li (2024, May). Simple Techniques for Enhancing Sentence Embeddings in Generative Language Models.

Zhang, P. (2023). Endogenous capital-augmenting R&D, intersectoral labor reallocation, and the movement of the labor share. *Journal of Economics*, 1–36.

Zhao, X., C. Wu, and D. Liu (2021). Comparative Analysis of the Life-Cycle Cost of Robot Substitution: A Case of Automobile Welding Production in China. *Symmetry 13*(2), 226.

A Appendix: Innovation in Human Tasks by Acemoglu and Restrepo (2019)

Acemoglu and Restrepo (2019) (henceforth referred to as AR) presents a tool for inferring automation and innovation in human tasks (henceforth, IHT). This tool utilizes a relatively small set of variables: labor compensation, employee count, value-added, wage, and investment price. The AR framework enables the inference of automation and IHT. Fundamentally, the AR framework operates under the assumption that if there is an observed *increase* in labor share, it must be attributed to IHT. Conversely, if there is a *decrease*, it is attributable to automation. This principle is clearly articulated in Figure 1 of their paper.

The online appendix of the AR paper elaborates on this framework. For ease of reference, we include it in our Appendix N. In this appendix, Term (AR4) represents the percentage change in labor share, which can be broken down into Terms (AR6) and (AR7). The former represents the percentage change in substitution effects, while the latter shows the percentage change in 'task contents.' A positive (negative) result in Term (AR7) is interpreted as indicative of IHT (automation). Given that the percentage change in substitution effects (Term AR6) is usually minimal, the percentage change in 'task contents' (Term AR7) virtually mirrors the percent change in labor share (Term AR4).

To summarize, AR's inference of automation and IHT is largely based on the percent change in labor share. However, using these inferred variables in our primary analysis presents a challenge due to the expected high correlation with labor share, which could lead to reverse causality. Furthermore, there is no certainty that the inferred variables accurately represent the real-world values of automation and IHT. Consequently, we require variables obtained through direct measurement.

B Appendix: Model

B.1 Households

The representative consumer consumes an aggregated continuum of final goods, with the mass of final goods assumed to be one for simplicity. It's also assumed that there is no disutility from the supply of labor. The utility function of the representative consumer takes the following form:

$$U = \left(\int_0^1 Y(i)^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}} \tag{13}$$

, where η represents the elasticity of substitution between final goods.

The representative consumer's budget constraint is as follows:

$$\int_0^1 P(i)Y(i)di = \int_0^1 \left(\int_{N-1}^{N} W_j l_j(i)dj + \int_{N-1}^{N} \psi m_j(i)dj + RK_i + \Pi_i \right) di \tag{14}$$

, where W_j, ψ, and R represent wage for labor conducting task j, robot price, and capital price, respectively.

B.2 Labor Share

A step-by-step process for this section is provided in Appendix C. We set an assumption related to robot and labor productivity for simple algebra in deriving the equilibrium in the model.

Assumption 1. $\psi < \frac{W_I}{\gamma_I}$

The above assumption implies that it is efficient to use a robot for task j below I. In other words, whenever firms have the technological capability to substitute labor with a robot, they would be inclined to do so. This is a reasonable assumption, especially considering that robot prices have significantly declined, while wages have seen a steady increase. Figure 6 illustrates these trends by depicting the 5-year growth rates of the respective prices.

Based on the Assumption 1 and by solving the firm's cost minimization problem, factor demands, the price for the aggregated task, and the marginal cost of firm i are derived as follows:

$$l_j(i) = 0, \;\; \text{if } j \leq I \tag{15}$$

$$l_j(i) = \gamma_j^{\zeta-1} \left(\frac{W_j}{P_T} \right)^{-\zeta} T(i), \;\; \text{if } j > I \tag{16}$$

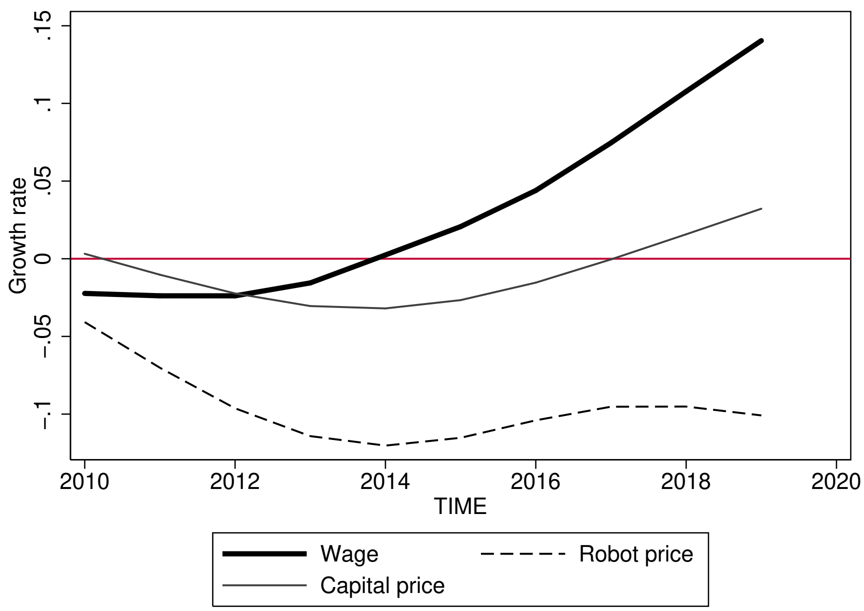
$$m_j(i) = \left(\frac{\psi}{P_T} \right)^{-\zeta} T(i), \;\; \text{if } j \leq I \tag{17}$$

$$m_j(i) = 0, \;\; \text{if } j > I \tag{18}$$

$$T(i) = \left(\frac{P_T}{MC(i)} \right)^{-\sigma} Y(i) \tag{19}$$

$$K(i) = \left(\frac{R}{MC(i)} \right)^{-\sigma} Y(i) \tag{20}$$

Figure 6: Prices in a 5-year growth rate



$$P_T = \left[(I - N + 1)\psi^{1-\zeta} + \int_I^N \left(\frac{W_j}{\gamma_j} \right)^{1-\zeta} dj \right]^{\frac{1}{1-\zeta}} \tag{21}$$

$$MC(i) = \left[P_T^{1-\sigma} + R^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{22}$$

$$W_j l_j(i) = \left(\frac{W_j}{\gamma_j} \right)^{1-\zeta} \cdot P_T^{\zeta} \cdot T_i \tag{23}$$

, where P_T and MC_i represent the price for the aggregated task and marginal cost of firm i, respectively.

C Appendix: Detailed Model Derivations

C.1 Environment

There is a representative household with utility function in Equation (24):

$$U = \left(\int_0^1 Y(k)^{\frac{\eta-1}{\eta}} dk \right)^{\frac{\eta}{\eta-1}}. \tag{24}$$

There are infinite number of identical firms i with production functions in Equation (27) and (28):

$$t_j(i) = m_j(i) + \gamma_j l_j(i) \ \text{ if } j \leq I \tag{25}$$
$$t_j(i) = \gamma_j l_j(i) \ \text{ if } j > I \tag{26}$$
$$T(i) = \left(\int_{N-1}^{N} t_j(i)^{\frac{\zeta-1}{\zeta}} dj \right)^{\frac{\zeta}{\zeta-1}} \tag{27}$$
$$Y(i) = \left(T(i)^{\frac{\sigma-1}{\sigma}} + K(i)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} . \tag{28}$$

By Assumption 1, Equation (25) simplifies to Equation (29). Without this assumption, the algebra becomes too complex to yield a closed-form solution. The implication of this assumption is that whenever robot operation is technically feasible, firms opt for robots over labor. This is because, according to Assumption 1, the cost of using a robot is lower than the cost of labor for unit of production.

$$t_j(i) = m_j(i) \ \text{ if } j \leq I \tag{29}$$

C.2 Step 1: derive P_T, and optimal inputs for robot* and labor*

We derive P_T, the price for an aggregated task, $T(i)$, by solving the cost minimization problem. We assume perfectly competitive market.

$$\min \text{cost(i) for } T(i) \text{ s.t. Equation(29), (26), and (27)}$$

$$\Rightarrow \min \int_{N-1}^{I} \psi m_j dj + \int_{I}^{N} w_j l_j dj \ \text{ s.t. } \left(\int_{N-1}^{I} m_j^{\frac{\zeta-1}{\zeta}} dj + \int_{I}^{N} (\gamma_j l_j)^{\frac{\zeta-1}{\zeta}} dj \right)^{\frac{\zeta}{\zeta-1}} = T(i)$$

$\Rightarrow$ This finds optimal inputs for robot* and labor* to produce T(i)

$\Rightarrow$ Specifically, letting T(i)=1 means the minimization solution is the price for T(i), P_T :

$$\Rightarrow P_T = \left[(I - N + 1)\psi^{1-\zeta} + \int_{I}^{N} \left(\frac{w_j}{\gamma_j} \right)^{1-\zeta} dj \right]^{\frac{1}{1-\zeta}} \tag{30}$$

C.3 Step 2: find optimal inputs for $T(i)$ and $K(i)$

Next, we find optimal inputs for $T(i)$ and $K(i)$ to produce $Y(i)$.

$$
\begin{aligned}
&\min \text{cost(i) for } Y(i) \text{ s.t. Equation(28)} \\
\Leftrightarrow &\min P_T \cdot T(i) + R \cdot K(i) \text{ s.t. Equation(28)} \\
\Rightarrow &\text{This finds optimal inputs for T(i)* and K(i)* to produce Y(i)} \\
\Rightarrow &\text{Specifically, the minimization solution is the minimum cost for producing } Y(i) \\
\Rightarrow &\begin{cases}
T(i)^* = Y(i) P_T^{-\sigma} \\
K(i)^* = Y(i) R^{-\sigma} \\
\text{Cost for } Y(i) = Y(i)\left[P_T^{1-\sigma} + R^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \\
\qquad\quad = Y(i) \times \text{AC} \\
\qquad\quad = Y(i)
\end{cases}
\end{aligned}
$$

We let $\left[P_T^{1-\sigma} + R^{1-\sigma}\right]^{\frac{1}{1-\sigma}} = 1$ as a numeraire. This numeraire significantly simplifies the algebraic complexity. Since we let AC$= 1$, MC is also one.

C.4 Step 3: find a demand function for $Y(i)$

Next, we find a demand function for $Y(i)$ by minimizing consumption cost.

$$
\begin{aligned}
&\min \text{cost for consumption s.t. Equation(24)} \\
\Leftrightarrow &\min \int_0^1 P(i)Y(i)di \text{ s.t. Equation(24)} \\
\Rightarrow &\text{Specifically, this yields a demand function for } Y(i) \\
\Leftrightarrow &Y(i) = \left(\frac{P(i)}{\mathbb{P}}\right)^{-\eta} \text{, where } \mathbb{P} \equiv \left[\int_0^1 P(i)^{1-\eta} di\right]^{\frac{1}{1-\eta}}
\end{aligned}
$$

C.5 Step 4: find firm(i)'s profit

The final goods market is the monopolistic competition that allows firms' positive profit. Until now, we know two things: (1) a demand function for $Y(i)$, and (2) the minimum cost for producing $Y(i)$. Firm's profit maximization problem yields:

$$
\begin{aligned}
P(i)^* &= \frac{\eta}{\eta - 1} \\
\Rightarrow \Pi(i) &= \frac{1}{\eta - 1} Y(i)^*
\end{aligned}
$$

Meanwhile, we naturally get optimal $Y(i)$ as below, but this is redundant for this paper.

$$Y(i)^* = \left(\frac{\eta}{(\eta-1)\mathbb{P}}\right)^{-\eta}, \text{ where } \mathbb{P} \equiv \left[\int_0^1 P(i)^{1-\eta} di\right]^{\frac{1}{1-\eta}}$$

C.6 Step 5: derive the labor cost for producing optimal $Y(i)$

In Step 1, we already found optimal inputs of $l_j(i)$ to produce $T(i)$. Therefore we can also know the optimal labor cost at task j for firm i to produce $T(i)$.

$$\begin{aligned} l_j(i)^* &= \left(\frac{W_j(i)}{\gamma_j P_T}\right)^{-\zeta} \gamma_j^{-1} T(i) \qquad (31) \\ \Rightarrow W_j(i) l_j(i)^* &= \left(\frac{W_j(i)}{\gamma_j}\right)^{1-\zeta} P_T^{\zeta} T(i) \end{aligned}$$

And we also derived optimal $T(i)$ while in Step 2: $T(i)^* = Y(i) P_T^{-\sigma}$. Plugging in this to the equation above,

$$W_j(i) l_j(i)^* = \left(\frac{W_j(i)}{\gamma_j}\right)^{1-\zeta} P_T^{\zeta-\sigma} Y(i)$$

Therefore, the optimal labor cost for firm i to produce $Y(i)$ by using every task from I to N is:

$$\begin{aligned} \int_I^N W_j(i) l_j(i)^* dj &= \int_I^N \left(\frac{W_j(i)}{\gamma_j}\right)^{1-\zeta} P_T^{\zeta-\sigma} Y(i) dj \\ &= \int_I^N \left(\frac{W_j(i)}{\gamma_j}\right)^{1-\zeta} dj \cdot P_T^{\zeta-\sigma} Y(i) \end{aligned}$$

C.7 Step 6: derive an expression for labor share

Until now, we have figured out (1) labor cost, (2) total cost, and (3) profit. Putting all together, we find labor share. Since we prefer not to focus on $\frac{\eta-1}{\eta}$, we move this term to the left-hand side.

$$
\begin{aligned}
S_L(i) &= \frac{\text{Labor cost(i)}}{\text{Total cost(i)} + \text{Profit(i)}} = \frac{\text{Labor cost(i)}}{Y(i) + \frac{1}{\eta-1}Y(i)} \\
&= \frac{\eta-1}{\eta}\frac{\text{Labor cost(i)}}{\text{Total cost(i)}} \\
\Leftrightarrow \frac{\eta}{\eta-1}S_L(i) &= \frac{\text{Labor cost(i)}}{\text{Total cost(i)}} \\
&\equiv S_L^f(i)
\end{aligned}
$$

After substituting the expressions for Labor cost(i) and Total cost(i) that we derived earlier, we finally construct a detailed expression for $S_L^f(i)$.

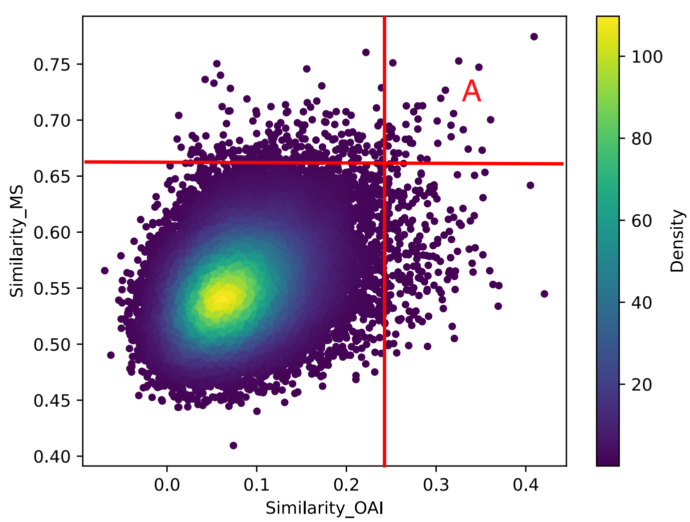
$$
\begin{aligned}
S_L^f(i) &= \frac{\text{Labor cost(i)}}{\text{Total cost(i)}} \\
&= \frac{\int_I^N W_j(i) l_j(i) dj}{Y(i)} \\
&= \frac{\int_I^N W_j(i) l_j(i) dj}{P_T T(i) + RK(i)} \\
&= \frac{\int_I^N \left(\frac{W_j(i)}{\gamma_j}\right)^{1-\zeta} dj \cdot P_T^{\zeta-\sigma} Y(i)}{P_T^{1-\sigma} Y(i) + R^{1-\sigma} Y(i)} \\
&= \frac{\int_I^N \left(\frac{W_j(i)}{\gamma_j}\right)^{1-\zeta} dj}{P_T^{1-\zeta}} \frac{P_T^{1-\sigma}}{P_T^{1-\sigma} + R^{1-\sigma}} \\
&\text{, where } P_T \equiv \left[(I - N + 1)\psi^{1-\zeta} + \int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj\right]^{\frac{1}{1-\zeta}}
\end{aligned}
$$

D Appendix: Patent Matching

We select the top 0.106 percent of similarity scores from the entire pool of matched pairs across years and sectors to ensure that the matches are highly relevant, excluding scores below this threshold. It is important to note that the numerator and denominator of 0.106% are not values representing the number of patents but rather the number of all possible many-to-many forced matches. Specifically, the total possible many-to-many matches amount to 2.890 billion. After eliminating irrelevant matchings, we utilize 3.063 million matchings. From the perspective of patents, the total number of available patents is 3,621,749. After the exclusion, the number of patents we utilize is 1,145,917, which constitutes 31.64% of all patents. Regarding occupations, the total number is 798, and after the exclusion, we utilize 791 occupations, indicating that seven occupations have zero matchings with patents.

As illustrated in Figure 7 using a few sample matches, an overwhelming majority of matchings between patents and occupations are likely irrelevant; thus, we need to exclude these (The x-axis represents the similarity scores generated by OpenAI's embedding vectors, while the y-axis shows scores from Microsoft. The figure depicts a classic two-dimensional scatter plot, with colors indicating kernel densities). If we do not implement this cutoff, the vast number of irrelevant matches —each score is below the cutoff— would dominate the sum of the similarity scores in each year, industry, and occupation, rendering this summed value meaningless.

Figure 7: Scatter Density Plot of Similarity Scores



The selection process is detailed as follows: After manual review, we determined that similarity scores below 0.66 (Microsoft embedding version) are less relevant for comparing descriptions between patents and occupations. Consequently, we retain only matches that exceed this cutoff value, which represents 0.766% of the total many-to-many matches. We apply the same criterion to the OpenAI embedding version, 'text-embedding-3-large'.[25] These cutoff values are depicted as red lines in Figure 7. We then select matches within the area labeled 'A' in this figure, comprising 0.106 percent of the

[25]The 'text-embedding-ada-002' model was a notable advancement in OpenAI's embedding technology, offering multi-language support and enhanced accuracy. Yet, newer models like 'text-embedding-3-small' and 'text-embedding-3-large' now outperform it, delivering better performance and cost-efficiency. The 'text-embedding-3-small' model shows substantial improvements in benchmarks like MIRACL (31.4% to 44.0%) and MTEB (61.0% to 62.3%), and costs 80% less. The larger 'text-embedding-3-

entire pool of matched pairs. Different embedding models may capture various aspects of semantic and syntactic relationships in text. By cross-checking cosine similarity scores across multiple models, we can mitigate the risk of relying on a single model that might exhibit specific biases or limitations. Mazandu et al. (2021) discusses an integrated platform that supports the calculation of semantic similarity scores using multiple applications.

Meanwhile, each patent includes IPC information, which can be categorized into industry sectors following the methodology outlined by Lybbert and Zolas (2014). Since the number of patents matched to occupations and their respective similarity scores vary by industry according to patent data, the weights applied also differ across industries. Consequently, the finalized IHT variable is industry-specific.

When using word-to-word embedding software, a cleaning procedure is necessary before the embedding step, involving the removal of prepositions and other stop words while retaining nouns and verbs. In contrast, sentence-to-sentence embedding software interprets entire sentences without this process, as removing stop words can disrupt the context and semantic meaning these models aim to capture (Khan et al., 2020).

To make the best use of the sentence embedding feature, we used lengthy text descriptions. For occupation descriptions, we included both titles and detailed explanations. For patent descriptions, we used detailed content from the body of the patent document, retaining the first 100 words which typically describe the purpose and essence of the patent. In contrast, some existing studies, such as Montobbio et al. (2021), utilize concise patent descriptions from IPC or CPC codes.

E Appendix: Adjusted Penetration of Robots

APR is defined as in Equation (32):

$$\mathrm{APR}_{i,(t5,t1)} \equiv \frac{M_{i,t5} - M_{i,t1}}{L_{i,2005}} - \frac{Y_{i,t5} - Y_{i,t1}}{Y_{i,t1}} \frac{M_{i,t1}}{L_{i,2005}} \tag{32}$$

$$= \Big(\frac{M_{i,t5} - M_{i,t1}}{M_{i,t1}} - \frac{Y_{i,t5} - Y_{i,t1}}{Y_{i,t1}}\Big) \frac{M_{i,t1}}{L_{i,2005}} \tag{33}$$

$$= \big(g_M - g_Y\big) \frac{M_{i,t1}}{L_{i,2005}} \tag{34}$$

, where i is the industry sector (country $\times$ industry in our case), and t5 is 5-year after t1. M is the number of robots (stock), L is the number of employees, Y is value-added

large' model boosts performance further, with MIRACL scores rising to 54.9% and MTEB to 64.6%. These models offer a deeper understanding of sentence meanings, making them preferable for tasks requiring text matching and understanding sentence semantics.

(in real terms).

Acemoglu and Restrepo (2020) employs APR as a proxy for $d(I - N + 1)$ primarily because the term dI encapsulates the theoretical concept of a 'pure direction of automation,' which is abstract and not directly observable in empirical settings. The observable growth rate of the number of robots is not a suitable proxy for dI since it reflects an equilibrium outcome in real-world scenarios. Given this, Acemoglu and Restrepo (2020) proposes APR to effectively serve as a proxy for $d(I - N + 1)$.

The second term in Equation (34), $-g_Y$, serves to measure the 'penetration' of robots. In other words, if the growth rate of robots exceeds that of value-added, they interpret this as a positive penetration. This penetration equates to $I - N + 1$ in their terminology, which represents the length between $N - 1$ and I. The inclusion of the second term, (34), $-g_Y$, in Equation (34) is necessary for the following reason: Suppose there is an economic boom. In such a scenario, the growth rate of robot adoption would likely surge, while $d(I - N + 1)$ remains unchanged. Therefore, they adjust the growth rate of robot adoption by subtracting the growth rate of value-added, g_Y.

The APR represents the 5-year growth rate of robots adjusted by labor input and the value-added within a given sector. Multiplication by $\frac{M_{i,t1}}{L_{i,2005}}$ is necessary as the raw number of robots does not adequately represent their definition of automation. Consider, for instance, that the IFR began collecting data in many countries starting in 2004. A change from 1 robot to 100 robots between 2004 and 2005 would represent a growth rate of 9900%, whereas an increase from 100 to 200 robots between 2005 and 2006 would only reflect a 100% growth rate. These rates are not useful because the number of machines increased by the same amount (100) in both cases. The term $\frac{M_{i,t1}}{L_{i,2005}}$ is introduced to adjust for this discrepancy. Suppose $L_{i,2005} = 100$. In 2005, $g_M \times \frac{M_{i,t1}}{L_{i,2005}}$ equals 99%, and in 2006, it amounts to 100%, which makes them comparable. The underlying idea is that the 5-year difference in the number of machines across countries and industries is not directly comparable; they needed to normalize it by dividing by the number of employees.[26]

F Appendix: Capital Price

In our paper, we utilize the replicated values for capital price from Karabarbounis and Neiman (2014) (hereinafter KN). To calculate this, we initially require the investment price, which the KLEMS data provides, including industry variations.

It's important to note that we don't directly observe the capital price, which represents the *usage* cost of one unit of capital. We do, however, observe the investment

[26] Instead of dividing by $L_{i,2005}$, dividing by 'quantity' would be more accurate, but it will not change the results significantly.

price, which signifies the *purchase* cost of one unit of capital. In accordance with the theory of investment by Jorgenson (1963), we can calculate the capital price as follows:

$$R_t = \xi_{t-1}(1 + i_t) - \xi_t(1 - \delta_t) \tag{35}$$

$$R_t = \xi_t \left(\frac{1}{\beta} - 1 + \delta \right) \tag{36}$$

In this Equation (35), R represents the capital price, ξ is the investment price, i is the interest rate, and δ is the depreciation rate. All values are expressed in real terms. This equation signifies that investors are indifferent between paying a *usage* cost for capital (R_t) and *purchasing* capital, paying interest, and then selling the depreciated capital at a later date.

To simplify Equation (35) into the form presented in Equation (36), we follow a specific process. This involves the assumption of a constant interest rate, i, and approximating $1+i$ as $\frac{1}{\beta}$. Equation (36), as employed by KN in their KLEMS version of the capital price variable, assumes a depreciation rate of 10%. This rate aligns closely with the 10.8% rate assumed by Stehrer et al. (2019), an official KLEMS document. Throughout this paper, we strictly adhere to the approach by KN.[27]

G Appendix: KLEMS Data and Capital Cost

G.1 KLEMS Data

Aside from the IFR dataset, the ONET dataset, and Robot Price, we will use data from KLEMS.[28] All nominal values are converted to real values through division by the chain-linked price index provided by KLEMS (VA_PI), following the methodology implemented by Karabarbounis and Neiman (2014).

KLEMS comes in two different versions: one follows national accounts, and the other follows growth accounts. The main difference between these versions is that the national accounts allow room for a markup greater than one, while the growth accounts do not. The latter assumes that the sum of labor cost and capital cost equals the value-added, implying that the markup is exactly one. As allowing for a markup is critical for our analysis, we use the national accounts when using KLEMS.

[27]It is important to note that KN employed a β value of 0.909 (corresponding to an interest rate, $i = 0.100$), reflecting the high real interest rates prevalent in the 1970s. In contrast, our study adopts a β of 0.988 (equivalent to $i = 0.012$), derived from averaging the real interest rates from 2005 to 2019 across ten countries. However, the specific value of β does not influence the regression outcomes in our analysis, as we focus on the growth rate of the capital price, which effectively cancels out the impact of β.

[28]KLEMS: EU level analysis of capital (K), labour (L), energy (E), materials (M) and service (S) inputs.

KLEMS shares similar characteristics with OECD STAN in terms of many national account variables at a country-industry-year level. Table 4 presents descriptive statistics. Predominantly, the values for OECD STAN and KLEMS are comparable, albeit not identical. In some instances, the values are in fact identical. This alignment is a result of collaborative projects aimed at fostering more consistent values between the two.

Table 4: Descriptive Statistics

Country	WL (labor comp)		RK (capital comp)		Value added		Labor Share	
	STAN	KLEMS	STAN	KLEMS	STAN	KLEMS	STAN	KLEMS
USA	867,789	851,834	292,456	308,662	1,647,140	1,593,719	52.85	53.60
DEU	366,787	366,806	104,117	104,034	569,189	570,196	64.67	64.57
SWE	256,507	256,540	115,040	124,370	502,728	502,728	51.17	51.18
DNK	219,076	226,496	199,337	220,713	410,478	426,533	55.33	54.87
ITA	140,568	140,568	57,107	54,924	253,368	253,353	55.60	55.60
FRA	135,093	135,098	52,379	41,244	226,181	226,181	59.74	59.74
GBR	110,603	109,347	26,230	25,535	171,778	170,498	64.45	64.19
AUT	28,106	29,959	9,427	12,090	51,011	54,254	55.22	55.31
FIN	17,100	17,979	7,512	7,204	33,112	34,848	51.91	51.85
PRT	11,537	12,897	3,166	3,166	20,575	23,030	56.06	55.99
Total	215,317	214,753	86,677	90,194	388,556	385,534	56.75	56.69

G.2 Capital Cost

The KLEMS data has one limitation: it lacks RK (rental cost for capital stock) and profit (operating surplus and mixed income). If either RK or Profit were available, we could deduce the other because Value-added is calculated as WL + RK + Profit. Regrettably, the absence of both presents a challenge. This issue is addressed by utilizing OECD STAN data.

In particular, the KLEMS dataset lacks RK. It does include I_GFCF (Investment in Gross Fixed Capital Formation) and K_GFCF (Capital Stock of Gross Fixed Capital Formation), but these do not provide the necessary RK information. I_GFCF represents the net investment in fixed assets —a flow metric indicating capital goods investment. K_GFCF, on the other hand, denotes the total value of all fixed assets available for production —a stock variable. Consequently, although RK can be estimated based on K_GFCF, this method lacks precision. This is because K_GFCF represents the purchase cost, not the rental cost. To convert the purchase cost into rental cost, the real interest rate and depreciation rate as shown in Equation (35) are required. Notably, the depreciation rate requires numerous assumptions, and we lack this information.

A pertinent question arises: why not use OECD STAN initially, instead of KLEMS? The response lies in the fact that OECD STAN does not contain R (capital price) data.

Therefore, we resort to using R obtained from KLEMS. However, integrating this with other data from OECD STAN, particularly wage variables, poses complications. Furthermore, STAN does not provide industry-specific Producer Price Index (PPI). To enhance the accuracy of our analysis, we prefer to use industry-specific PPI, specifically the VA_PI variable from KLEMS.

Hence, an alternative approach is to employ RK from OECD STAN. This is feasible because the value-added and WL (labor compensation) figures are nearly identical in both STAN and KLEMS datasets (as illustrated in Figures 9 in Appendix M). Consequently, it is highly probable that RK, along with operating surplus and mixed income, are consistent across both KLEMS and STAN. Therefore, in this paper, we assume that the markups in KLEMS and STAN are identical, denoted by $\frac{\text{Value-added}}{\text{WL+RK}}$. Based on this assumption, we are able to recover RK for KLEMS as below:

$$\frac{\text{Value-added}_{\text{STAN}}}{\text{WL}_{\text{STAN}} + \text{RK}_{\text{STAN}}} = \frac{\text{Value-added}_{\text{KLEMS}}}{\text{WL}_{\text{KLEMS}} + \mathbf{RK_{KLEMS}}}.$$

H Appendix: Estimation of S_M^T

Denote Ψ, M, W, and L as robot price, number of robots, wage, and employment, respectively. Then S_M^T can be expressed as follows:

$$\begin{aligned} S_M^T &= \frac{\Psi M}{\Psi M + WL} \\ &= \frac{1}{1 + \frac{WL}{\Psi M}} \\ &= \frac{1}{1 + \left(\frac{M}{L}\right)^{-1} \frac{W}{\Psi}} \end{aligned}$$

Unfortunately, the International Federation of Robotics (IFR) provided robot prices in the form of an average unit price until 2009 and discontinued this practice thereafter. Access to robot price information prior to 2009 is also restricted for those who have purchased IFR data after this point. Nonetheless, Fernandez-Macias et al. (2021) offers a comprehensive method to approximate the missing price information from the IFR dataset. Specifically, they provide values for M/L as well as Ψ. We supplement these data with wage information from the OECD STAN database to complete the S_M^T value in the equation above.

It is important to note that the equipment cost for robots is estimated to constitute around 33.04% of the total robot costs[29], covering elements like operation, training, soft-

[29] $33.04\% = 35.73\% \times (1 - 0.075)$, where 0.075 represents taxes, transactions, and after-sales fees. The cost share of robot equipment accounts for 35.73% of the total cost for using robots, as estimated by Zhao et al. (2021).

ware, maintenance, and disposal (Zhao et al., 2021). The figures provided by Fernandez-Macias et al. (2021) pertain only to equipment cost. Therefore, we have accounted for this information accordingly.

H.1 An Alternative Approach to Estimating the S_M^T

Let's assume labor cost to be 100 without loss of generality. According to KLEMS data, the rental cost for OMach is recorded as 13.595. But it's important to note that OMach encompasses not just robots but also a range of other items, including equipment, machinery, engines, and turbines (Stehrer et al., 2019; Gouma and Timmer, 2013). Therefore, the challenge is to determine the share of robots within the broader category of OMach. The most reliable approach we can consider involves utilizing UN Comtrade data, which offers information about import and export values by detailed commodity categories. By calculating the total export values of commodities corresponding to OMach,[30] and separately calculating the total export values of HS Code 8479 (which pertains to robots),[31] we find that the ratio between these values is 13.595 : 0.71. In brief, the ratio between labor cost, OMach cost, and robot cost is 100 : 13.595 : 0.71.

The equipment cost for robots is estimated to be around 33.04% of the total robot costs (Zhao et al., 2021), and the UN Comtrade estimate of 0.71 corresponds to the equipment cost. Therefore, the total cost of the robot amounts to $0.71/0.33 = 2.149$. Hence, S_M^T is estimated to be 2.104%.[32]

I Appendix: Estimation of σ and ζ

Given that $S_K^f > 0$ and the coefficient for $d\ln R$ is negative, we can infer that $\sigma < 1$. Further, by substituting the value $S_K^f = 0.494$ that we obtained from the data, we calculate $\sigma = 0.611$, as illustrated in Equation (37). We conduct a Wald test on the null hypothesis that $\sigma = 0$ and find that it can be rejected at the 0.05 significance level. The confidence interval for σ is $(0.477, 0.746)$. Consequently, we can conclude with confidence that σ lies within the range of 0 to 1.

$$-\underbrace{S_K^f}_{0.494}(1-\sigma) = \underbrace{\alpha_5}_{-0.19205} \tag{37}$$

$$\Rightarrow \sigma = 1 + \frac{\alpha_5}{S_K^f} \tag{Sigma}$$

[30] HS Classification 84 excluding 8401, 8402, 8403, 8404, 8405, 8429, 8440, 8443, 8470, 8471, and 8472.

[31] Machinery and mechanical appliances; having individual functions, n.e.c. in this chapter.

[32] $2.104\% = \frac{2.149}{2.149+100}$

The derivation of the value for ζ proceeds as follows. From Equation (9), utilizing coefficients α_3 and α_5, we arrive at Equation (Zeta).

$$\zeta = 1 - \frac{\alpha_3 + \alpha_5 S_L^T}{1 - S_L^T} \tag{Zeta}$$

As demonstrated earlier in Section 5.3, we estimate S_L^T to be 0.972. Upon substituting $S_L^T = 0.972$ into Equation (Zeta), we obtain an estimate for ζ of 2.374. We then conduct a Wald test on the null hypothesis that $\zeta = 0$ and find it can be rejected at the 0.05 significance level. Specifically, the confidence interval is from 1.017 to 3.730. Consequently, we can conclude with confidence that ζ lies within the range of this interval.

J Appendix: Direct and Indirect Effects for Automation and Innovation in Human Tasks

Automation: The term Ⓐ in Equation (9) denotes the direct effect of automation on labor share, which is negative. Concurrently, the term Ⓑ×Ⓒ captures the indirect effect. Specifically, Ⓒ is negative under Assumption 1, irrespective of the sign of ζ. This indicates that the price of the aggregated task, denoted by P_T, falls when robots take over tasks previously performed by humans. This change in P_T is then scaled by the factor $-(1-\zeta) + S_K^f(1-\sigma)$, which represents the partial derivative of labor share with respect to the aggregated task price. Therefore, the sign of the indirect effect on labor share hinges critically on the sign of $-(1-\zeta) + S_K^f(1-\sigma)$, which we have estimated to be positive. In summary, given that Ⓑ > 0 and Ⓒ < 0, the indirect effect of automation on labor share is also negative, serving to amplify its direct impact.

Innovation in Human Tasks: The term Ⓓ in Equation (9) denotes the direct effect of IHT on labor share, which is positive. Concurrently, the term Ⓑ×Ⓔ captures the indirect effect. Specifically, Ⓔ is positive under Assumption 1, with the proof provided the next subsection. This indicates that the price of the aggregated task, denoted by P_T, rises when there is innovation in human tasks. Since Ⓑ> 0, the indirect effect of IHT on labor share is also positive, serving to amplify its direct impact.

J.1 Proof of $E > 0$

Here, we explain why Ⓔ is positive. To do this, we rewrite Equation (9) as below:

$$d \ln S_L^f = \ldots d \ln \gamma + \alpha_1 dI + \alpha_2 dN + \ldots d \ln \mathbb{W} + \ldots d \ln R + \ldots d \ln \psi.$$

We can rearrange $\alpha_1 dI + \alpha_2 dN$ as follows:

$$\begin{aligned}
&\alpha_1 dI + \alpha_2 dN \\
=&\alpha_1 dI - \alpha_1 dN + \alpha_1 dN + \alpha_2 dN \\
=&\alpha_1 d(I - N + 1) + (\alpha_1 + \alpha_2) dN \\
=&\alpha_1 d(I - N + 1) + \beta_2 dN
\end{aligned}$$

, where β_2 is

$$\beta_2 = \underbrace{\left(S_N^L - S_I^L\right) \frac{1}{1-\zeta}}_{\text{F}} \underbrace{\left[S_M^T(1-\zeta) + S_L^T S_K^f (1-\sigma)\right]}_{\text{G}} \tag{38}$$

We can estimate this β_2 by a regression. We perform the identical regression in Column (1) of Table 2 except that we use APR instead of IRB. The regression result is provided in Section 7 Table 8, in which $\beta_2 = 0.00019 > 0$. The sign of G in Equation (38) is positive because the robot cost share, denoted as S_M^T, is a very small value, specifically 0.028. Given that β_2 is positive, F in Equation (38) is also positive. Since S_N^L and S_I^L are defined as $\frac{\left(\frac{W_N}{\gamma_N}\right)^{1-\zeta}}{\int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj}$ and $\frac{\left(\frac{W_I}{\gamma_I}\right)^{1-\zeta}}{\int_I^N \left(\frac{W_j}{\gamma_j}\right)^{1-\zeta} dj}$, respectively, the sign of $\left(S_N^L - S_I^L\right)\frac{1}{1-\zeta}$ is the same as that of $\left[\left(\frac{W_N}{\gamma_N}\right)^{1-\zeta} - \left(\frac{W_I}{\gamma_I}\right)^{1-\zeta}\right]\frac{1}{1-\zeta}$, which is a positive value. Assumption 1 asserts that $\psi < \frac{W_I}{\gamma_I}$. This assumption is reasonable, given the observed decline in robot prices and the corresponding increase in wages (Figure 6). Combining this assumption with $\left[\left(\frac{W_N}{\gamma_N}\right)^{1-\zeta} - \left(\frac{W_I}{\gamma_I}\right)^{1-\zeta}\right]\frac{1}{1-\zeta}$ establishes that the sign of $\frac{1}{1-\zeta}\left[-\psi^{1-\zeta} + \left(\frac{W_N}{\gamma_N}\right)^{1-\zeta}\right]$ is positive. In summary, E in Equation (9) is positive.

K Appendix: Estimation of the Elasticity of Substitution between Labor and Non-robot Capital

The condition $\sigma < 1$ indirectly confirms that capital and labor are gross complementary, a result that aligns with the findings reported by Glover and Short (2020). Conversely, this result contradicts the hypothesis of gross substitutability ($\sigma > 1$) posited by Karabarbounis and Neiman (2014) (henceforth KN). We clarify that the term σ in our general equilibrium model does not align exactly with the definition of σ in the work of KN as well as Glover and Short (2020). The divergence stems from our model's distinction between robots and non-robot capital. Specifically, in our model, σ represents the elasticity of substitution between 'non-robot capital' and 'aggregated tasks', where the latter encompasses both robot and labor inputs.

Hence, in this subsection, we introduce the elasticity of substitution between labor and non-robot capital, denoted by μ, a measure that closely aligns with the findings of both KN and Glover and Short (2020). The solution for μ is given in Equation (39), and its derivation can be found in Appendix L.

$$
\begin{aligned}
\mu \equiv & \frac{d\left(\frac{L}{K}\right)}{d\left(\frac{R}{\mathbb{W}}\right)} \frac{\frac{R}{\mathbb{W}}}{\frac{L}{K}} \quad \text{, where} \\
d\left(\frac{L}{K}\right) = & \left(\frac{\mathbb{W}_1}{R_1}\right)^{-\sigma} \left[\frac{S_M^T}{1-S_M^T}\left(\frac{\mathbb{W}_0}{\mathbb{W}_1}\right)^{1-\zeta} + 1\right]^{\frac{\zeta-\sigma}{1-\zeta}} - \left(\frac{\mathbb{W}_0}{R_0}\right)^{-\sigma} \left[\frac{S_M^T}{1-S_M^T} + 1\right]^{\frac{\zeta-\sigma}{1-\zeta}} \\
\frac{L}{K} = & \left(\frac{\mathbb{W}_0}{R_0}\right)^{-\sigma} \left[\frac{S_M^T}{1-S_M^T} + 1\right]^{\frac{\zeta-\sigma}{1-\zeta}} \\
\Rightarrow \ & \mu = \sigma \ \text{ if } S_M^T = 0.
\end{aligned}
\tag{39}
$$

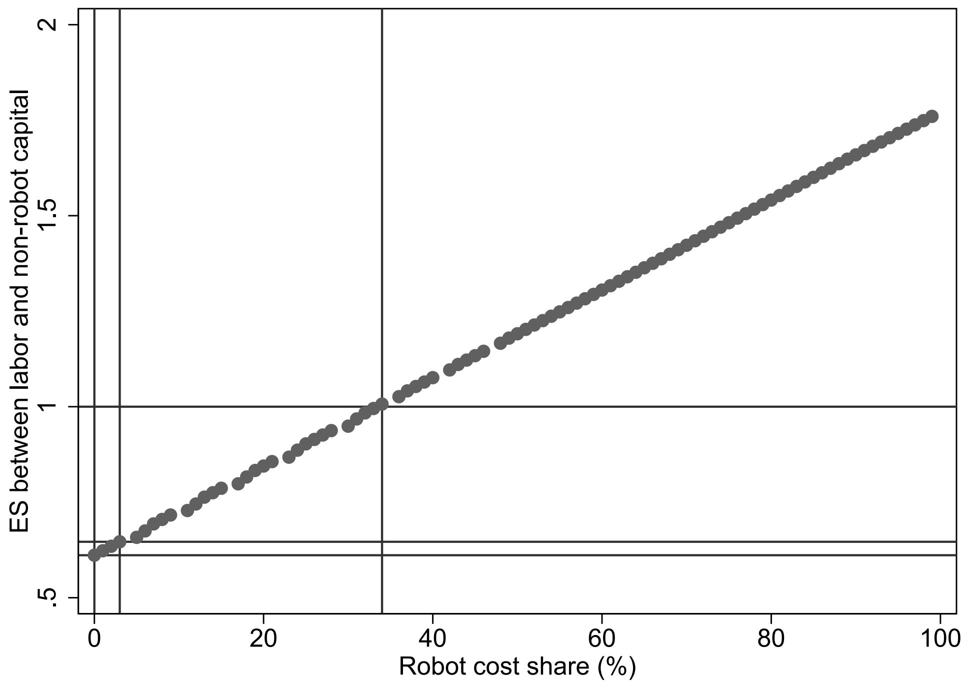
Differentiating Equation (39) is infeasible. However, we can employ numerical approximation to estimate μ. We use actual $\mathbb{W}$ and R values from the dataset (all possible combinations of these), along with $\sigma = 0.611$. We introduce small random variations to each $\mathbb{W}$ and R and consider scenarios where $|\Delta \frac{R}{\mathbb{W}}|$ is approximately 0.01. These values are then plugged into Equation (39) to obtain an approximated μ.

Panel (a) of Figure 8 displays the approximation results. When S_M^T is zero, we find that $\mu = \sigma = 0.611$. This stage indicates a complete absence of robot tasks, with all tasks being performed by labor. When $S_M^T = 2.813\%$, which corresponds to our estimate presented in Section 5.3, we obtain $\mu = 0.644$. Even when we assume $S_M^T = 10\%$, the divergence from σ is minimal, reaching at most $\mu = 0.722$. It is only at $S_M^T = 34\%$ that μ exceeds one. Consequently, we argue that in the context of the KN model, the elasticity of substitution between labor and non-robot capital closely approximates σ. Our analysis suggests that μ ranges between 0.611 and 0.722, supporting the idea of a gross complementary relationship between the two. In the future, as automated robots assume a greater share of tasks, the elasticity of substitution between labor and non-robot capital may approach, or even exceed, one —particularly if the robot share surpasses 34%. However, making accurate predictions about this trend necessitates more comprehensive research.

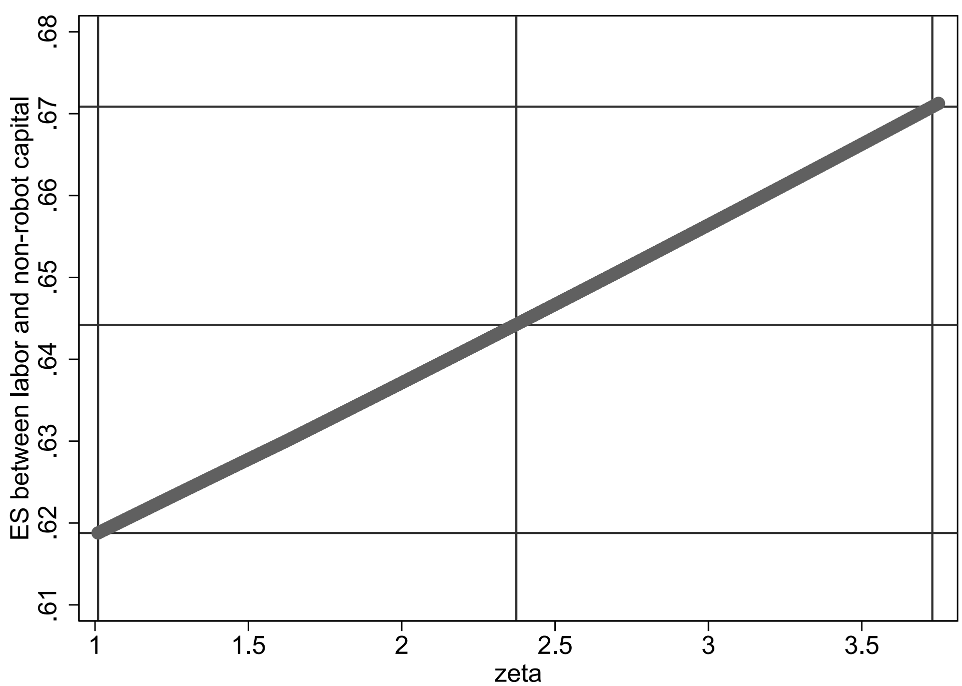
The above estimation of μ is contingent upon the value of $\zeta = 2.374$, which is our point estimate as derived in Section 5.4. However, the confidence interval for ζ varies: it spans from 1.017 to 3.730. To demonstrate the robustness of our μ estimate, we examine its sensitivity across a wide range of ζ values. This analysis is presented in Panel (b) of Figure 8. Within the ζ range of 1.017 to 3.730, μ varies between 0.619 and 0.671, confirming the robustness of our μ estimation.

Figure 8: Elasticity of Substitution between Labor and Non-robot Capital

(a) Fixing ζ to be 2.525; Moving S_M^T



(b) Fixing S_M^T to be 2.813%; Moving ζ



Recent research underscores the importance of quantifying this elasticity of substitution between labor and capital, as highlighted by Martinez (2018), Oberfield and Raval (2021), and Zhang (2023). Many studies report an elasticity less than one, endorsing the concept of gross complementarity. However, Piketty and Zucman (2014) suggest the potential for gross substitutability. They observed an escalating capital-output ratio and argued that this trend could consistently account for the declining labor share if the elasticity of substitution between labor and capital exceeds one —a claim our estimates do not corroborate.

Our finding also does not support the hypothesis proposed by Karabarbounis and Neiman (2014), who argue that the falling price of capital accounts for half of the recent decline in labor share. For their argument to hold, the elasticity of substitution between labor and capital must be greater than one (gross substitutes). They directly measured the correlation between the trend of capital price and labor share without using instrumental variables.

In contrast, Glover and Short (2020) reached a different conclusion, that of gross complements, by using cross-country variation with instrumental variables. They argue that correcting for bias is critical when estimating the correlation between the capital price and labor share. Our paper addresses omitted variable bias using a control function approach. We regress automation, the emergence of new tasks, wages, and robot price, along with capital price, on labor share, believing that this approach corrects for omitted variable bias. Our study supports Glover and Short (2020).

L Appendix: Derivation of μ

Let μ denote the elasticity of substitution between labor and non-robot capital. The concept of elasticity of substitution formally defines μ as follows:

$$\mu \equiv \frac{d\left(\frac{L}{K}\right)}{d\left(\frac{R}{\mathbb{W}}\right)} \frac{\frac{R}{\mathbb{W}}}{\frac{L}{K}}. \tag{40}$$

To proceed, we must express L and K in terms of W and R, respectively. Equation (31), derived in Appendix C.6, provides the formulation for L as follows:

$$\begin{aligned} l_j(i)^* &= \left(\frac{W_j(i)}{\gamma_j P_T}\right)^{-\zeta} \gamma_j^{-1} T(i) \\ \Rightarrow L &= \int_I^N l_j(i)^* dj \\ &= \int_I^N \left(\frac{W_j(i)}{\gamma_j P_T}\right)^{-\zeta} \gamma_j^{-1} T(i) dj. \end{aligned} \tag{41}$$

We introduce a parameter β_j to serve as a weight for the wage distribution corresponding to each worker, indexed by j. Utilizing β_j enables us to establish a representative measure for wages, $\mathbb{W}$.

$$W_j \equiv \beta_j \mathbb{W} \tag{42}$$

Consequently, Equation (41) can be restructured to yield Equation (43). To streamline the notation, we define $A = \int_I^N \gamma_j^{\zeta-1}\beta_j^{-\zeta} dj$.

$$L = \int_I^N \gamma_j^{\zeta-1}\beta_j^{-\zeta} dj \ \cdot T(i)\Big(\frac{\mathbb{W}}{P_T}\Big)^{-\zeta} \tag{43}$$

$$= A \cdot T(i)\Big(\frac{\mathbb{W}}{P_T}\Big)^{-\zeta} \tag{44}$$

We have derived $T(i)$ in Appendix C.3 and P_T in Appendix C.2. For the sake of clarity, we restate these formulations here:

$$T(i) = Y(i)P_T^{-\sigma}$$

$$P_T = \left[(I-N+1)\psi^{1-\zeta} + \int_I^N \left(\frac{w_j}{\gamma_j}\right)^{1-\zeta} dj\right]^{\frac{1}{1-\zeta}}$$

By substituting $T(i)$ and P_T into Equation (44),

$$\begin{aligned} L =& A \cdot Y(i)P_T^{-\sigma}\Big(\frac{\mathbb{W}}{P_T}\Big)^{-\zeta} \\ =& A \cdot Y(i)P_T^{\zeta-\sigma}\mathbb{W}^{-\zeta} \\ =& A \cdot Y(i)\left[(I-N+1)\psi^{1-\zeta} + \int_I^N \left(\frac{w_j}{\gamma_j}\right)^{1-\zeta} dj\right]^{\frac{\zeta-\sigma}{1-\zeta}} \mathbb{W}^{-\zeta}. \end{aligned}$$

$(I-N+1)\psi^{1-\zeta}$ and $\int_I^N \left(\frac{w_j}{\gamma_j}\right)^{1-\zeta} dj$ correspond to the cost share of robots and human labor, respectively. Consequently, we can reformulate these expressions as follows:

$$(I-N+1)\psi^{1-\zeta} \equiv S_M^T$$

$$\int_I^N \left(\frac{w_j}{\gamma_j}\right)^{1-\zeta} dj \equiv S_L^T$$

Therefore, L can be reformulated as follows:

$$L = A \cdot Y(i)\Big[S_M^T + S_L^T\Big]^{\frac{\zeta-\sigma}{1-\zeta}}\mathbb{W}^{-\zeta}$$

$$= A \cdot Y(i)\left[\frac{S_M^T}{S_L^T} + 1\right]^{\frac{\zeta-\sigma}{1-\zeta}}\mathbb{W}^{-\zeta} \tag{45}$$

We derived the optimal value of K in Appendix C.3, given by $K = Y(i)R^{-\sigma}$. Consequently, we complete our derivation of $\frac{L}{K}$ as follows:

$$\begin{aligned}\frac{L}{K} =& \frac{A \cdot Y(i)\left[\frac{S_M^T}{S_L^T}+1\right]^{\frac{\zeta-\sigma}{1-\zeta}}\mathbb{W}^{-\zeta}}{Y(i)R^{-\sigma}}\\ =& \frac{A \cdot \left[\frac{S_M^T}{S_L^T}+1\right]^{\frac{\zeta-\sigma}{1-\zeta}}\mathbb{W}^{-\zeta}}{R^{-\sigma}}\end{aligned}$$

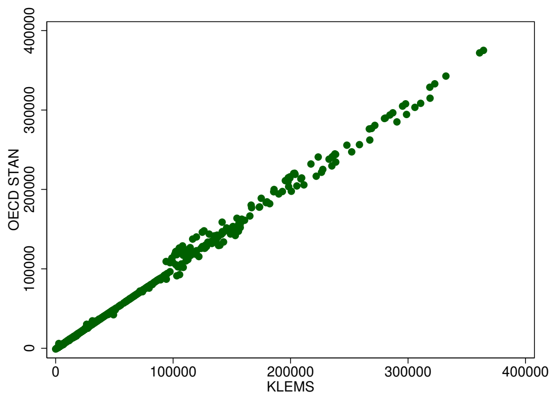
Thus, the expression for $d\left(\frac{L}{K}\right)/\frac{L}{K}$ is given below. This concludes our derivation of μ.

$$\frac{d\left(\frac{L}{K}\right)}{\frac{L}{K}} = \frac{\left(\frac{\mathbb{W}_1}{R_1}\right)^{-\sigma}\left[\frac{S_M^T}{1-S_M^T}\left(\frac{\mathbb{W}_0}{\mathbb{W}_1}\right)^{1-\zeta}+1\right]^{\frac{\zeta-\sigma}{1-\zeta}} - \left(\frac{\mathbb{W}_0}{R_0}\right)^{-\sigma}\left[\frac{S_M^T}{1-S_M^T}+1\right]^{\frac{\zeta-\sigma}{1-\zeta}}}{\left(\frac{\mathbb{W}_0}{R_0}\right)^{-\sigma}\left[\frac{S_M^T}{1-S_M^T}+1\right]^{\frac{\zeta-\sigma}{1-\zeta}}}$$

M Appendix: Tables and Figures

Figure 9: Values by Country, Sector, and Year

(a) Value-added



(b) Labor Compensation[33]

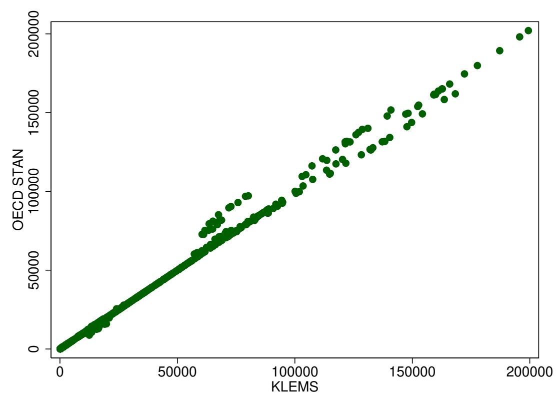


Figure 10: KN's Capital Prices

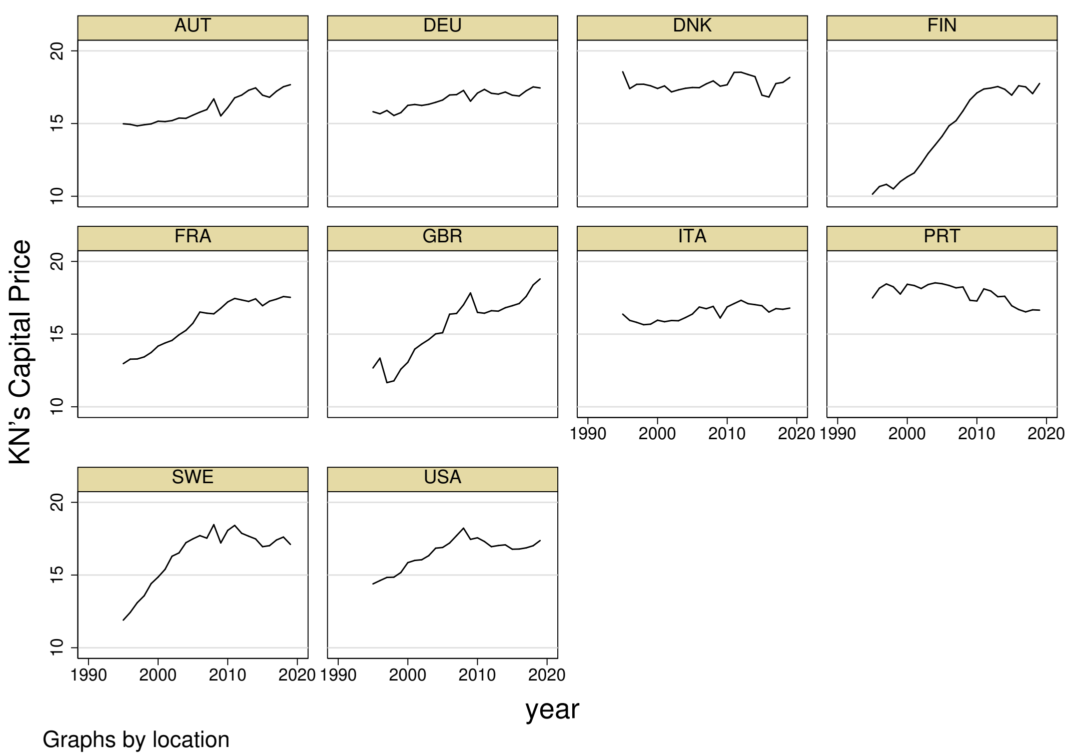


Table 5: Regressions (unweighted)

	With labor productivity			Without labor productivity	
	(1)	(2)	(3)	(4)	(5)
Restrction 1	No	Yes	Yes	No	Yes
Restrction 2	No	No	Yes	No	NA
α_1: IRB	-0.090***	-0.093***	-0.119***	-0.073***	-0.058**
	(0.023)	(0.025)	(0.021)	(0.021)	(0.025)
α_2: IHT	0.110***	0.112***	0.142***	0.096***	0.081***
	(0.025)	(0.027)	(0.021)	(0.024)	(0.028)
α_3: gr_labor price	14.771***	16.341***	4.314***	12.098***	11.802***
	(3.637)	(3.604)	(1.657)	(3.067)	(3.440)
α_4: gr_robot price	0.383	0.306	1.196	0.552	0.617
	(0.984)	(0.997)	(1.042)	(0.987)	(1.024)
α_5: gr_non robot capital price	-19.651***	-16.648***	-5.510***	-19.184***	-12.419***
	(3.422)	(3.349)	(1.687)	(3.416)	(3.261)
α_6: gr_labor productivity	-2.712*	-4.452***	-4.314***		
	(1.426)	(1.426)	(1.657)		
N	998	998	998	998	998
R^2	0.654			0.650	

The coefficients and the standard errors have been multiplied by 100 for better readability.
Standard errors in parenthesis are clustered by country and sector.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Regressions (wage weighted)

	With labor productivity			Without labor productivity	
	(1)	(2)	(3)	(4)	(5)
Restrction 1	No	Yes	Yes	No	Yes
Restrction 2	No	No	Yes	No	NA
α_1: IRB	-0.091***	-0.095***	-0.122***	-0.075***	-0.061**
	(0.024)	(0.026)	(0.021)	(0.022)	(0.026)
α_2: IHT	0.113***	0.116***	0.142***	0.098***	0.085***
	(0.024)	(0.026)	(0.021)	(0.022)	(0.026)
α_3: gr_labor price	14.872***	16.382***	4.258***	12.255***	11.963***
	(3.571)	(3.557)	(1.640)	(3.023)	(3.381)
α_4: gr_robot price	0.053	0.001	0.857	0.185	0.251
	(0.943)	(0.949)	(1.013)	(0.949)	(0.972)
α_5: gr_non robot capital price	-19.307***	-16.383***	-5.116***	-18.815***	-12.214***
	(3.395)	(3.314)	(1.695)	(3.385)	(3.200)
α_6: gr_labor productivity	-2.671*	-4.358***	-4.258***		
	(1.375)	(1.369)	(1.640)		
N	998	998	998	998	998
R^2	0.656			0.652	

The coefficients and the standard errors have been multiplied by 100 for better readability.
Standard errors in parenthesis are clustered by country and sector.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Regressions (variance unadjusted IHT)

	With labor productivity			Without labor productivity	
	(1)	(2)	(3)	(4)	(5)
Restrction 1	No	Yes	Yes	No	Yes
Restrction 2	No	No	Yes	No	NA
α_1: IRB	-0.089***	-0.093***	-0.118***	-0.071***	-0.056**
	(0.022)	(0.024)	(0.020)	(0.020)	(0.024)
α_2: IHT	0.155***	0.156***	0.205***	0.138***	0.119***
	(0.036)	(0.036)	(0.033)	(0.035)	(0.037)
α_3: gr_labor price	14.801***	16.321***	4.574***	11.873***	11.595***
	(3.512)	(3.512)	(1.599)	(2.996)	(3.367)
α_4: gr_robot price	0.007	-0.042	0.732	0.162	0.243
	(0.940)	(0.944)	(1.008)	(0.953)	(0.983)
α_5: gr_non robot capital price	-19.205***	-16.280***	-5.306***	-18.668***	-11.838***
	(3.353)	(3.267)	(1.679)	(3.360)	(3.161)
α_6: gr_labor productivity	-2.982**	-4.665***	-4.574***		
	(1.395)	(1.397)	(1.599)		
N	998	998	998	998	998
R^2	0.655			0.650	

The coefficients and the standard errors have been multiplied by 100 for better readability.
Standard errors in parenthesis are clustered by country and sector.
$^* p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$

Table 8: Regressions (using APR)

	With labor productivity			Without labor productivity	
	(1)	(2)	(3)	(4)	(5)
Restrction 1	No	Yes	Yes	No	Yes
Restrction 2	No	No	Yes	No	NA
β_1: APR	-0.089***	-0.093***	-0.118***	-0.071***	-0.056**
	(0.022)	(0.024)	(0.020)	(0.020)	(0.024)
β_2: IHT	0.019**	0.018**	0.025***	0.020**	0.018**
	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)
β_3: gr_labor price	14.801***	16.321***	4.574***	11.873***	11.595***
	(3.512)	(3.512)	(1.599)	(2.996)	(3.367)
β_4: gr_robot price	0.007	-0.042	0.732	0.162	0.243
	(0.940)	(0.944)	(1.008)	(0.953)	(0.983)
β_5: gr_non robot capital price	-19.205***	-16.280***	-5.306***	-18.668***	-11.838***
	(3.353)	(3.267)	(1.679)	(3.360)	(3.161)
β_6: gr_labor productivity	-2.982**	-4.665***	-4.574***		
	(1.395)	(1.397)	(1.599)		
N	998	998	998	998	998
R^2	0.655			0.650	

The coefficients and the standard errors have been multiplied by 100 for better readability.
Standard errors in parenthesis are clustered by country and sector.
$^* p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$

Table 9: Estimations for ζ, σ, and $-(1-\zeta)+S_K^f(1-\sigma)$

	ζ	σ	$-(1-\zeta)+S_K^f(1-\sigma)$
Main Specification (Table 2)	2.374 (1.017, 3.730)	0.611 (0.477,0.746)	1.566 (0.197, 2.934)
IHT unweighted (Table 5)	2.538 (1.081, 3.995)	0.602 (0.465, 0.740)	1.735 (0.266, 3.203)
IHT wage weighted (Table 6)	2.383 (0.980, 3.787)	0.609 (0.473, 0.746)	1.576 (0.161, 2.992)
IHT variance adjusted (Table 7)	2.374 (1.017, 3.730)	0.611 (0.477, 0.746)	1.566 (0.197, 2.934)
Using APR (Table 8)	2.374 (1.017, 3.730)	0.611 (0.477, 0.746)	1.566 (0.197, 2.934)

Figure 11: Labor shares (unweighted IHT)

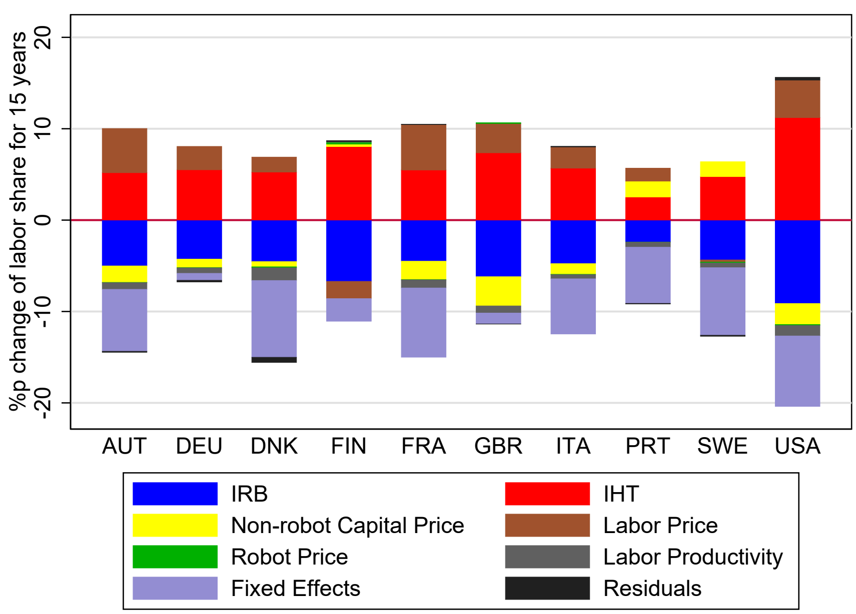


Figure 12: Labor shares (wage weighted IHT)

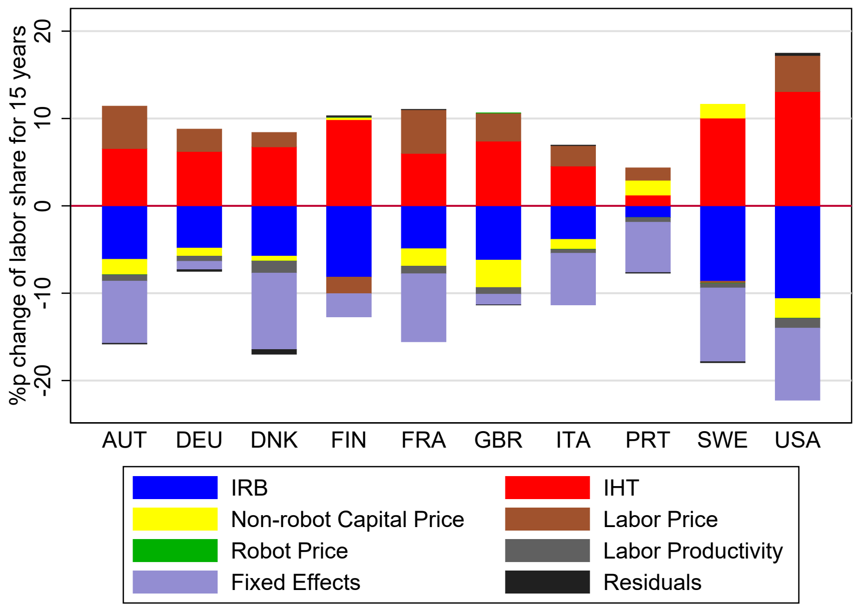


Figure 13: Labor shares (variance unadjusted IHT)

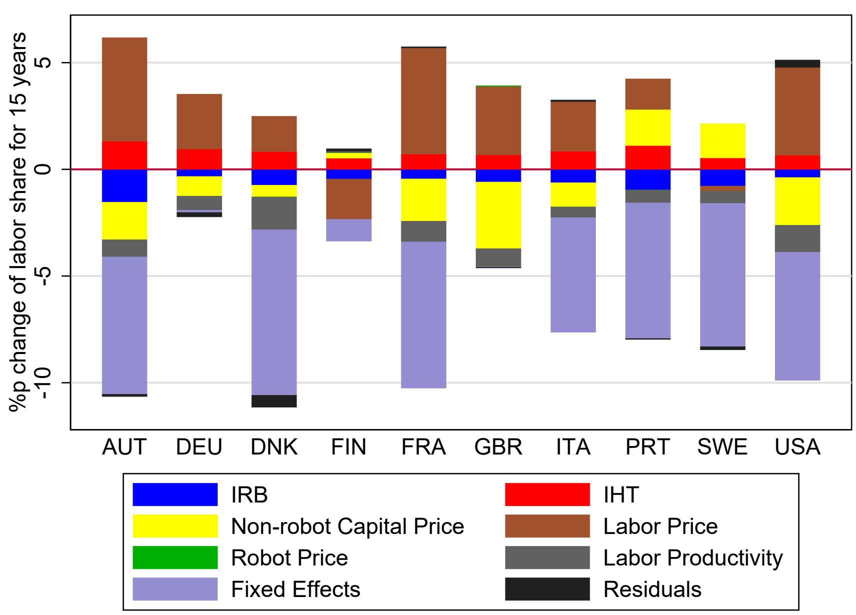
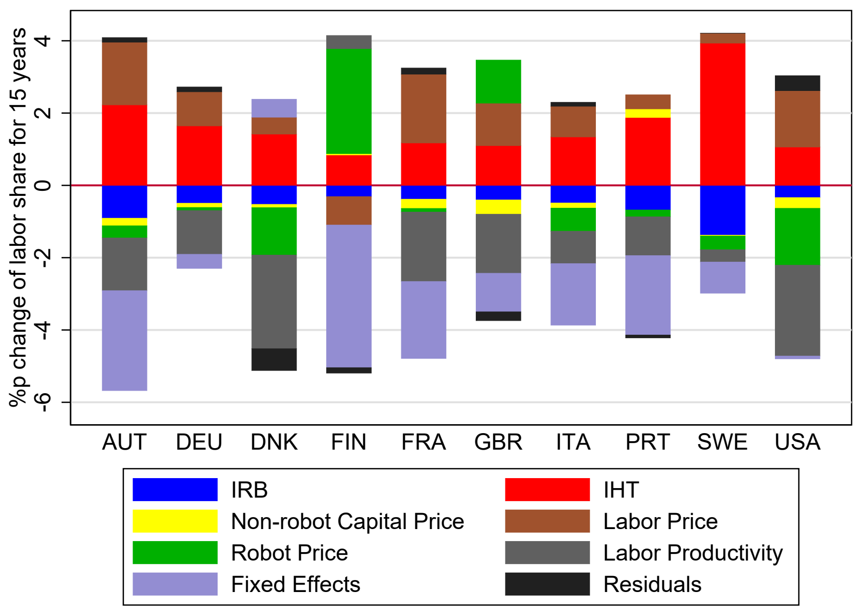


Figure 14: Labor shares (based on Steady-State Analysis)



N Appendix: Acemoglu and Restrepo (2019)

Let me first introduce their notations in Table 10.

Table 10

Notation	Meaning
i	Industry sector
P_i	The price of the goods produced by sector i
Y_i	Output (value added) of sector i
$Y = \sum_i P_i Y_i$	Total value added (GDP) in the economy
$\chi_i = \frac{P_i Y_i}{Y} = \frac{P_i Y_i}{\sum_i P_i Y_i} = \frac{\text{GDP}_i}{\text{GDP}}$	The share of sector i's GDP
W_i	Wage per worker in sector i
L_i	Number of workers in sector i
$W_i L_i$	Total wage bill in sector i
$WL = \sum_i W_i L_i$	Total wage bill in the economy
$\ell_i = \frac{W_i L_i}{WL}$	The share of the wage bill in sector i
$s_i^L = \frac{W_i L_i}{P_i Y_i} = \frac{\text{Total wage bill}_i}{\text{GDP}_i}$	The labor share in sector i
$s^L = \frac{WL}{Y} = \frac{\text{Total wage bill}}{\text{GDP}}$	The labor share in the economy
$\Gamma_i = \Gamma(N_i, I_i)$	The task content of production with regards to labor in sector i
γ_i^L	The comparative advantage schedules for labor in sector i
γ_i^K	The comparative advantage schedules for captial in sector i

The decomposition starts from the percent change in the wage bill normalized by population (Equation (AR1)). Since $\ln\left(\frac{W_t L_t}{N_t}\right)$ can be expressed as $\ln\left(Y_t \sum_i \chi_{it} s_{it}^L\right)$, Equation (AR1) can be decomposed as Equation (AR2);

$$
\begin{aligned}
&\ln\left(\frac{W_t L_t}{N_t}\right) - \ln\left(\frac{W_{t0} L_{t0}}{N_{t0}}\right) && \text{(AR1)}\\
&= \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right) && \text{(AR2)}\\
&+ \ln\left(\sum_i \chi_{it} s_{it}^L\right) - \ln\left(\sum_i \chi_{it0} s_{it0}^L\right)\\
&= \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right)\\
&+ \ln\left(\sum_i \chi_{it} s_{it}^L\right) - \ln\left(\sum_i \chi_{it0} s_{it}^L\right)\\
&+ \ln\left(\sum_i \chi_{it0} s_{it}^L\right) - \ln\left(\sum_i \chi_{it0} s_{it0}^L\right)\\
&\approx \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right)\\
&+ \sum_i \frac{s_{it}^L}{\sum_j \chi_{jt0} s_{jt}^L}\left(\chi_{it} - \chi_{it0}\right)\\
&+ \ln\left(\sum_i \chi_{it0} s_{it}^L\right) - \ln\left(\sum_i \chi_{it0} s_{it0}^L\right)\\
&\approx \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right) && \text{(AR3)}\\
&+ \sum_i \frac{s_{it}^L}{\sum_j \chi_{jt0} s_{jt}^L}\left(\chi_{it} - \chi_{it0}\right)\\
&+ \sum_i \ell_{it0}(\ln s_{it}^L - \ln s_{it0}^L) && \text{(AR4)}
\end{aligned}
$$

The first-order Taylor expansion of Term (AR4) yields Terms (AR6) and (AR7); Denote $(1-\sigma)(1-s_{it0}^L)\left(\ln \frac{W_{it}}{W_{it0}} - \ln \frac{R_{it}}{R_{it0}} - g_{i,t0,t}^A\right)$ as Substitution$_{i,t0,t}$, we can rewrite Equation (AR5) as (AR8); Denote $\left(\ln s_{it}^L - \ln s_{it0}^L\right) -$Substitution$_{i,t0,t}$ as ChangeTaskContent$_{i,t0,t}$, we can rewrite Equation (AR8) as (AR9).

$$
\begin{aligned}
&\approx \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right) && \text{(AR5)}\\
&+ \sum_i \frac{s^L_{it}}{\sum_j \chi_{jt0} s^L_{jt}} \left(\chi_{it} - \chi_{it0}\right)\\
&+ \sum_i \ell_{it0} \Bigg[(1-\sigma)(1-s^L_{it0}) \Bigg(\ln\frac{W_{it}}{W_{it0}} - \ln\frac{R_{it}}{R_{it0}} - g^A_{i,t0,t} \Bigg) && \text{(AR6)}\\
&\qquad + \frac{1-s^L_{it0}}{1-\Gamma_{it0}} (\ln\Gamma_{it} - \ln\Gamma_{it0}) \Bigg] && \text{(AR7)}\\
&\approx \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right)\\
&+ \sum_i \frac{s^L_{it}}{\sum_j \chi_{jt0} s^L_{jt}} \left(\chi_{it} - \chi_{it0}\right)\\
&+ \sum_i \ell_{it0} \Bigg[\text{Substitution}_{i,t0,t}\\
&\qquad + \frac{1-s^L_{it0}}{1-\Gamma_{it0}} (\ln\Gamma_{it} - \ln\Gamma_{it0}) \Bigg]\\
&\approx \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right) && \text{(AR8)}\\
&+ \sum_i \frac{s^L_{it}}{\sum_j \chi_{jt0} s^L_{jt}} \left(\chi_{it} - \chi_{it0}\right)\\
&+ \sum_i \ell_{it0} \Bigg[\text{Substitution}_{i,t0,t}\\
&\qquad + \left(\ln s^L_{it} - \ln s^L_{it0} \right) - \text{Substitution}_{i,t0,t} \Bigg]\\
&\approx \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right) && \text{(AR9)}\\
&+ \sum_i \frac{s^L_{it}}{\sum_j \chi_{jt0} s^L_{jt}} \left(\chi_{it} - \chi_{it0}\right)\\
&+ \sum_i \ell_{it0} \Bigg[\text{Substitution}_{i,t0,t}\\
&\qquad + \text{ChangeTaskContent}_{i,t0,t} \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&\approx \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right) \\
&+ \sum_i \frac{s^L_{it}}{\sum_j \chi_{jt0} s^L_{jt}} \left(\chi_{it} - \chi_{it0}\right) \\
&+ \text{Substitution}_{t0,t} \\
&+ \sum_i \ell_{it0} \Big[\text{ChangeTaskContent}_{i,t0,t}\Big]
\end{aligned}
$$

$\sum_i \ell_{it0}[\text{ChangeTaskContent}_{i,t0,t}]$ can be decomposed again into Equation (AR10), assuming that over five-year windows, an industry engages in either automation or the creation of new tasks but not in both activities.

$$
\text{Displacement}_{t-1,t} = \sum_{i \in \mathcal{I}} \ell_{i,t0} \min\left\{0, \frac{1}{5} \sum_{\gamma=t-2}^{t+2} \text{ChangeTaskContent}_{i,\gamma-1,\gamma}\right\} \tag{AR10}
$$

$$
\text{Reinstatement}_{t-1,t} = \sum_{i \in \mathcal{I}} \ell_{i,t0} \max\left\{0, \frac{1}{5} \sum_{\gamma=t-2}^{t+2} \text{ChangeTaskContent}_{i,\gamma-1,\gamma}\right\}
$$

To sum up, starting from Equation (AR1), it can be decomposed into 1) productivity, 2) composition, 3) substitution, 4) displacement, and 5) reinstatement effects.

$$
\begin{aligned}
&\ln\left(\frac{W_t L_t}{N_t}\right) - \ln\left(\frac{W_{t0} L_{t0}}{N_{t0}}\right) && \text{[Wage bill per capita]} \\
&\approx \ln\left(\frac{Y_t}{N_t}\right) - \ln\left(\frac{Y_{t0}}{N_{t0}}\right) && \text{[Productivity effect]} \\
&+ \sum_i \frac{s^L_{it}}{\sum_j \chi_{jt0} s^L_{jt}} \left(\chi_{it} - \chi_{it0}\right) && \text{[Composition effect]} \\
&+ \text{Substitution}_{t0,t} && \text{[Substitution effect]} \\
&+ \text{Displacement}_{t0,t} && \text{[Displacement effect (Automation)]} \\
&+ \text{Reinstatement}_{t0,t} && \text{[Reinstatement effect (New tasks)]}
\end{aligned} \tag{AR11}
$$