

Graetz, Georg; Feng, Andy

**Conference Paper**

## Rise of the Machines: The Effects of Labor-Saving Innovations on Jobs and Wages

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik  
- Session: Job Polarization, No. D14-V2

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Graetz, Georg; Feng, Andy (2014) : Rise of the Machines: The Effects of Labor-Saving Innovations on Jobs and Wages, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Job Polarization, No. D14-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100401>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Rise of the Machines: The Effects of Labor-Saving Innovations on Jobs and Wages\*

ANDY FENG<sup>†</sup> & GEORG GRAETZ<sup>‡</sup>

This version: December 2, 2013

First version: August 17, 2013

## JOB MARKET PAPER

### Abstract

Job polarization—the rise in employment shares of high and low skill jobs at the expense of middle skill jobs—occurred in the US not just recently, but also in the late nineteenth and early twentieth centuries. We argue that in each case polarization resulted from increased automation, and provide a theoretical explanation. In our model, firms deciding whether to employ machines or workers in a given task weigh the cost of using machines, which is increasing in the *complexity* (in an engineering sense) of the task, against the cost of employing workers, which is increasing in *training time* required by the task. Insights from artificial intelligence and robotics suggest that some tasks do not require training regardless of complexity, while in other tasks training is required and increases in complexity. In equilibrium, firms are more likely to automate a task that requires training, holding complexity constant. We assume that more-skilled workers learn faster, and thus it is middle skill workers who have a comparative advantage in tasks that are most likely to be automated when machine design costs fall. In addition to explaining job polarization, our model makes sense of observed patterns of automation and accounts for a set of novel stylized facts about occupational training requirements.

**Keywords:** job polarization, technical change, automation, wage inequality, training

**JEL codes:** E25, J23, J31, M53, O33

---

\*We thank Francesco Caselli, Guy Michaels, Alan Manning, Luis Garicano, John Van Reenen, Steve Pischke, Gianmarco Ottaviano, Thomas Sampson, and Michael Boehm for valuable comments and suggestions.

<sup>†</sup>London School of Economics and Singapore Ministry of Trade and Industry

<sup>‡</sup>Corresponding author; London School of Economics and Center for Economic Performance, Houghton Street, London WC2A 2AE, UK, g.graetz@lse.ac.uk.

# 1 Introduction

A growing empirical literature argues that recent technical change has led to polarization of labor markets in the US and Europe:<sup>1</sup> Employment in middle skill occupations has grown relatively more slowly than in high and low skill occupations since the 1980s, and similarly wages have grown faster at the top and bottom of the distribution than in the middle. While factors such as offshoring and changes in labor market institutions may have contributed to these developments, a consensus has emerged that identifies technical change as the main culprit. Modern information and communication technology (ICT) appears to substitute for workers in middle skill jobs, while complementing labor in high and low skill jobs, thus causing the observed reallocation of employment.<sup>2</sup>

Recent work has also documented *historical* instances of job polarization. Katz and Margo (2013) show that from 1850 to 1880, US manufacturing witnessed a relative decline in middle skill jobs (artisans) at the expense of high skill jobs (non-production workers) and low skill jobs (operators), concurrent with the increased adoption of steam power. Gray (2013) finds that electrification in the US during the first half of the 20th century led to a fall in demand for dexterity-intensive tasks performed by middle skill workers, relative to manual and clerical tasks performed by low and high skill workers, respectively.<sup>3</sup>

The innovations that preceded the three instances of job polarization have in common that they facilitated a more wide-ranging automation of tasks. The steam engine was instrumental in the increased mechanization of manufacturing because it provided a more reliable power source than water, and it allowed production to be located away from water, thus lowering transportation costs (Atack, Bateman, and Weiss 1980). Electricity facilitated automation because electric motors could be arranged much more flexibly than steam engines (Boff 1967). ICT allowed for the automation of cognitive tasks as well as improved control of physical production processes. Thus, the steam engine, electricity and ICT all triggered waves of labor-saving innovations. Current advances in artificial intelligence and robotics are likely to further boost automation, raising the question how job tasks, the distribution of employment across occupations, and the wage distribution will change as a result.<sup>4</sup>

In this paper, we argue that labor market polarization is inherent to labor-saving technologies. Our explanation is based on two insights. First, when technologies are available that can carry out a wide range of tasks autonomously, the allocation of workers and machines to tasks will be determined by comparative advantage (Simon 1960). Second, what is complex from an engineering point of view is not necessarily difficult to humans in the sense of requiring a lot of training (Moravec 1988). There are tasks that are easy to *any* worker but building a machine capable of performing them may be costly if not impossible; occupations such as waiters, taxi drivers,

---

<sup>1</sup>Job polarization has first been documented for the US by Autor, Katz, and Kearney (2006), for the UK by Goos and Manning (2007), and for other European economies by Goos, Manning, and Salomons (2009).

<sup>2</sup>See Autor, Levy, and Murnane (2003), Michaels, Natraj, and Van Reenen (forthcoming), and Goos, Manning, and Salomons (2011) for evidence favoring the technological explanation.

<sup>3</sup>Previously, Goldin and Katz (1998) presented evidence suggesting that electrification in the US was an instance of skill-biased technical change, although their empirical work focussed on a high-vs.-low-skill dichotomy.

<sup>4</sup>We provide a list of examples for recent progress in these areas in Section 2.5.

or housekeepers are intensive in the use of vision, movement, and communication, which are complex functions from an engineering point of view (Moravec 1988). On the other hand, a task like bookkeeping requires knowledge of arithmetic which takes humans years to learn, but which is trivial from an engineering perspective. We demonstrate how the principle of comparative advantage and the distinction between engineering complexity and difficulty in the sense of training time, combine to generate an equilibrium in which workers in the middle of the skill distribution are at the greatest risk of being replaced by machines.

Existing literature explains job polarization as caused by new technologies that are most suited for use in tasks initially performed by middle skill workers. Autor, Katz, and Kearney (2006) assume that ICT substitutes for routine tasks performed by middle skill workers, while complementing manual and abstract tasks that in equilibrium are carried out by low and high skill workers, respectively. Acemoglu and Autor (2011) present a model in which machines may replace any type of labor, but for the model to generate job polarization they must assume that machines replace middle skill workers. Our model imposes no such restrictions. Firms may employ labor of any type or machines to complete any given task, and will choose the factor that minimizes costs. There is no assumption about which part of the skill distribution will be most affected by increased automation—instead, this is endogenous to the attributes of factors and their interplay with the characteristics of tasks.

Labor-replacing technical change in our model refers to an exogenous fall in the cost of making machines, resulting from innovations that facilitate the automation of a wide range of tasks. Workers differ continuously by skill—higher skilled workers learn faster. Tasks are differentiated by *complexity* and by whether workers require training. Complexity refers to the intrinsic difficulty of a task, for instance the amount of information processing required, or the degrees of freedom and dexterity necessary to carry out a certain physical action. Complexity can be seen as a measure of ‘engineering difficulty’: the more complex a task, the larger an expenditure is required to build a machine capable of performing it. Over time, falling design costs make it cheaper to build machines for any task, and the complexity-cost gradient becomes flatter.

The model features a set of tasks of varying complexity for which workers do not need any training as they draw on capabilities that are innate or have been acquired at an early age—we call these *innate ability* tasks. Crushing rocks to produce gravel is an innate ability task of low complexity, while waiting tables is an innate ability task of high complexity, due to the amount of physical coordination required as well as the need for communicating in natural language. In the remaining tasks, workers cannot draw on an endowment of abilities and hence knowledge must be acquired. We call these tasks *training-intensive*, and we assume that training time increases in complexity in these tasks. Bookkeeping is a training-intensive task of low complexity: humans must acquire knowledge of arithmetic and the rules of accounting, but arithmetic and accounting rules are easily codified and require little processing power when executed by machines. Lawyering, on the other hand, is both training-intensive and highly complex: humans need to acquire knowledge of the legal system from scratch, and complexity arises from the need to apply a large set of general rules to specific cases, where rules and cases are described in different kinds of language. Both bookkeeping and lawyering are training-intensive tasks, but training takes

longer for lawyers, because their task is more complex.

The distinction between *innate ability* tasks and *training-intensive* tasks is critical for explaining why middle skill workers are most affected by increasing automation. Firms deciding whether to employ machines or workers in a given task weigh the cost of using machines, which is increasing in the complexity of the task, against the cost of employing workers, which is increasing in training time required by the task. As training time increases in complexity only in training-intensive tasks, firms are more likely to automate a training-intensive task than an innate ability one, holding complexity constant. Since higher skilled workers learn faster, workers at the bottom of the skill distribution have a comparative advantage (CA) in innate ability tasks. Workers at the top have a CA in highly complex training-intensive tasks, and workers in the middle have a CA in training-intensive tasks of intermediate complexity. In addition, we assume that design costs are such that machines' CA is in less complex tasks. Hence, the tasks most likely to be newly automated when machine design costs fall are those performed by middle-skill workers. It is in these tasks that the incentives for automation are strongest, due to the need to train workers.

Our model is able to explain job polarization as well as the hollowing out of the wage structure: given that middle-skill workers are most affected by machine replacement, they experience the strongest downward pressure on wages. The model provides a precise mechanism explaining labor market polarization, suggesting that the ICT revolution has caused labor market polarization because it has facilitated a more wide-ranging automation of tasks. This implies that the model is also suited to explain the historical instances of job polarization that followed the introduction of steam power and electricity in manufacturing.

Our model helps to explain recent patterns of automation. There are job tasks that are currently not automated despite the fact that it would be feasible to do so. For example, fast food preparation takes place in a controlled environment and involves a limited number of simple steps. It is thus not surprising that technology exists that automates this process (Melendez 2013). However, fast food jobs do not currently appear to be at risk of being automated. Moreover, in manufacturing some tasks are automated that are arguably of higher complexity (Davidson 2012). Given that fast food preparation requires very little training, in particular compared to non-trivial assembly tasks in manufacturing, this pattern of automation choices is exactly what our model would predict.

In addition to explaining labor market polarization and patterns of task automation, our theory delivers several novel predictions about trends in occupational training requirements. In the model, training requirements are higher in more complex training-intensive tasks. We measure training requirements in the US at the occupation level, using the Dictionary of Occupational Titles (DOT) combined with the 1971 April CPS, and the O\*NET database combined with the 2008 ACS.<sup>5</sup>

We find empirical support for the model's prediction of a polarization in training requirements, i.e. an increase in the employment shares of jobs requiring minimal and very high levels of training

---

<sup>5</sup>In these data, training requirements are measured as the time it takes the typical worker to become proficient in her job. This may include any occupation-specific knowledge acquired prior to entering the labor market—see Section 5.2.

between 1971 and 2007. Furthermore, we show that occupations that initially had intermediate training requirements experienced a fall in training requirements. The model provides a ready explanation: new technologies induced firms to automate the subset of tasks in a given occupation which required intermediate training by workers.<sup>6</sup> We also find that employment growth was less in occupations that experienced larger decreases in training requirements, as should be the case if automation causes training requirements to fall. Finally, we show that consistent with the model, returns to schooling have evolved differentially by occupational training requirements: returns have decreased in occupations that require very little training, but have increased in the remaining occupations.

The plan of the paper is as follows. The following subsection discusses related literature. Section 2 presents and solves the model. Section 3 discusses comparative statics, in particular how job assignment and the wage distribution change as a response to increased automation. We also present comparative statics for a change in skill supplies. Section 4 presents two extensions to the model: endogenous capital accumulation and a fixed cost of technology adoption. Section 5 confronts the model’s prediction with existing empirical evidence and takes novel implications of the model to the data. Section 6 concludes. All proofs are contained in Appendix B.

## 1.1 Related Literature

This paper is related to a recent literature featuring task-based models of exogenous job polarization. Autor, Levy, and Murnane (2003, henceforth ALM) categorize tasks as routine and non-routine. They call a task routine “if it can be accomplished by machines following explicit programmed rules” (ibid., p.1283). In contrast, non-routine tasks are “tasks for which rules are not sufficiently well understood to be specified in computer code and executed by machines” (ibid.). Extending the framework of ALM, Autor, Katz, and Kearney (2006) and Autor and Dorn (2013) allow non-routine tasks to be either low or high skill intensive.

We believe that our framework offers several advantages over ALM’s. First, it is not context-dependent. Machine capabilities constantly expand, so we prefer to avoid a task construct that depends on the current state of technology.<sup>7</sup> Complexity in our model is an objective, time-invariant measure of a task’s intrinsic difficulty. Second, the notion that worker training may be uncorrelated with a task’s complexity does not feature prominently in ALM’s framework. Finally, ALM’s framework implicitly leaves firms little choice to automate a given task, as routine tasks are assumed to be automated, and non-routine tasks are not. Our framework instead allows us to endogenize this choice. A task like fast food preparation might be considered “routine” in ALM’s framework, so that the non-automation of this task poses a puzzle. As noted above, our model is consistent with this example.

Acemoglu and Autor (2011) allow machines to replace labor in any task in principle, but assume that machines perform tasks initially performed by middle skill workers to make the

---

<sup>6</sup>All our results on changes in training requirements are robust to controlling for changes in mean years of schooling.

<sup>7</sup>To give an example, Levy and Murnane (2004) consider taking a left-turn on a busy road a nonroutine task unlikely to be automated in the foreseeable future. But less than a decade later, the driverless car has become a reality.

model consistent with job polarization. Because their task index does not have an empirical interpretation independently of factor assignment,<sup>8</sup> the model does not impose any restrictions on the data and thus could be consistent with arbitrary patterns of labor-replacing technical change, but without explaining them.<sup>9</sup>

We build on the literature on labor-saving innovations. Zeira (1998) presents a model in which economic development is characterized by the adoption of technologies that reduce labor requirements relative to capital requirements. Over time, an increasing number of tasks can be produced by new, more capital-intensive technologies. In an extreme example which is closely related to our paper, new technologies only use capital, while old ones only use labor. We extend this type of setting by explicitly modeling the characteristics of tasks and thus the direction of technical change, as well as by allowing for heterogeneous workers. Holmes and Mitchell (2008) present a model of firm organization where the problem of matching workers and machines to tasks is solved at the firm level. Their model admits a discrete set of worker types and they do not consider technical change.

The paper is related to a theoretical literature that uses assignment models to investigate the effects of technical change on the role of workers in the production process and on the wage distribution. Garicano and Rossi-Hansberg (2006) analyze how hierarchical organizations are affected by declines in communication and knowledge acquisition costs as caused by the ICT revolution. They match a continuum of skills to a continuum of tasks and find that lower communication costs lead to falling inequality among one group (workers) but to rising inequality among another group (managers). Thus, their model explains wage polarization. We abstract from issues related to firm organization and focus instead on the labor-saving aspect of ICT and other innovations. Another strand of papers analyzes the matching of workers with technologies of different vintages. Wage inequality results for instance when workers must acquire vintage-specific skills (Chari and Hopenhayn 1991) or machines are indivisible (Jovanovic 1998). Furthermore, skill or unskill bias of technical change can arise when new technologies require different learning investments than old ones, and when learning costs are a function of skill (Caselli 1999). We abstract from the issue of workers having to learn how to operate new technologies and focus instead on the problem of assigning workers and machines to tasks.

On the methodological side our paper is in the tradition of Ricardian theories of international trade, combining aspects of Dornbusch, Fischer, and Samuelson (1977) and Costinot and Vogel (2010). While these papers characterize equilibrium allocations *given* factor endowments and productivity levels, our focus is on endogenizing productivity differences, using modeling techniques similar to those of Costinot (2009). We shed light on the *sources* of comparative advantage between differently-skilled workers and machines.

---

<sup>8</sup>The task index in their model indicates whether high skill workers have a comparative advantage over middle skill workers in a given task, whether middle skill workers have a comparative advantage over low skill workers etc.

<sup>9</sup>Autor (2013) argues that “capital typically takes over tasks formerly performed by labor; simultaneously, workers are typically assigned novel tasks before they are automated.” One difficulty with this argument is that it does not help to explain why middle skill workers should be more affected by automation than low skill ones. A second difficulty is that there are many tasks that have existed for hundreds of years and are still not being automated (especially in low-skill services), while production of new products and services often involves tasks that workers never performed.

## 2 The Model

### 2.1 Overview

The model has one period that we interpret as a worker's lifetime.<sup>10</sup> There is a unique final good that is produced using a continuum of intermediate inputs, or *tasks*. These tasks are performed by workers of different skill levels and machines. Crucially, all factors of production are perfect substitutes at the task level. Although this may seem a strong assumption, the loss of generality is not substantial provided all tasks are essential in producing the final good, a condition that we shall maintain throughout. In fact, when tasks are imperfect substitutes in producing the final good, factors of production will appear to be imperfect substitutes in the aggregate.

Labor services as well as the economy's capital stock are supplied inelastically and all firms are perfectly competitive. Intermediate firms hire workers or capital to produce task output that is then sold to final good firms. Intermediate firms must train workers (except in innate ability tasks), and must transform generic capital into task-specific machines in order for these factors to be capable of performing tasks.

Technologies for worker training and machine design are public knowledge. Training and design expenditures are determined by task characteristics alone. We make this assumption for ease of exposition. In Appendix A we present a more general model that allows firms to choose training and design expenditures, thus determining factors' productivity endogenously. That model is based on an explicit characterization of the production process following Garicano (2000). All our results apply to the more general model as well.

### 2.2 Machine Design and Worker Training

Tasks are differentiated by *complexity*, denoted by  $\sigma \in [\underline{\sigma}, \bar{\sigma}]$ , and by whether workers require training to complete the task, indicated by  $\tau \in \{0, 1\}$ . The expenditure required to convert one unit of capital into a machine capable of performing a task of complexity  $\sigma$  is given by  $c_K \sigma$ . Complexity  $\sigma$  is the task-specific component of design expenditure. It is a fundamental, time-invariant property of tasks. The design cost  $c_K$  is constant across tasks and falls over time as better technologies become available, leading to a flatter complexity-cost gradient.<sup>11</sup>

For workers, the complexity of a task does not necessarily affect the amount of training required. In particular, no training is required if completion of a task relies solely on functions that all workers are endowed with or have acquired at early age, regardless of the complexity of the task. This is true in the case of innate ability tasks ( $\tau = 0$ ). Workers cannot rely on such endowments in the case of training-intensive tasks ( $\tau = 1$ ). Training requirements do increase with complexity in training-intensive tasks: To become capable of performing the training-intensive task of complexity  $\sigma$ , a worker of type  $s$  requires an amount of training  $\sigma/s$ . Higher skilled workers face a flatter complexity-training gradient in training-intensive tasks.

---

<sup>10</sup>We discuss a dynamic (multi-period) version of the model in Section 4.1.

<sup>11</sup>Strictly speaking,  $c_K$  is the design cost per unit of capital, per unit of the complexity measure.



### 2.3 Task Production

A machine produces  $A_K$  units of task output, regardless of the task's complexity. Hence,  $A_K$  can be viewed as task-neutral machine productivity. Since some of the hired capital is lost in machine design, a unit of capital produces an amount of output equal to  $A_K(1 - c_K\sigma)$ .<sup>12</sup>

All workers have a unit endowment of time, and produce one unit of task output if they are able to spend all their time in production. In other words, we normalize all workers' task-neutral productivity to one. Hence, workers of any type produce one unit of task output in innate ability tasks ( $\tau = 0$ ). Taking into account training time, a worker of type  $s$  produces  $1 - \sigma/s$  units of task output in training-intensive tasks ( $\tau = 1$ ) of complexity  $\sigma$ .

The notation  $s_K \equiv 1/c_K$ , 'machine skill', will turn out to be more convenient. Let worker type range from  $\underline{s} > 0$  to  $\bar{s}$ . To make the model interesting, we assume throughout that  $s_K, \underline{s} \geq \bar{\sigma}$ , so that machines and all worker types produce non-negative output in any task.

Let  $k_\tau(\sigma)$  denote the amount of capital used to produce task  $(\sigma, \tau)$  and similarly let  $n_\tau(s, \sigma)$  be the amount of type- $s$  labor. Given the task-specific productivity schedules for machines

$$\alpha^K(s_K, \sigma) = 1 - \sigma/s_K \tag{1}$$

and labor

$$\alpha^N(s, \sigma, \tau) = \begin{cases} 1 & \text{if } \tau = 0 \\ 1 - \sigma/s & \text{if } \tau = 1, \end{cases} \tag{2}$$

task output  $y$  can be written as

$$y_\tau(\sigma) = A_K \alpha^K(s_K, \sigma) k_\tau(\sigma) + \int_{\underline{s}}^{\bar{s}} \alpha^N(s, \sigma, \tau) n_\tau(s, \sigma) ds. \tag{3}$$

### 2.4 Final Good Production and Market Clearing

Let  $Y$  denote the output of the final good. For tractability, we use a Cobb-Douglas production function,

$$\log Y = \frac{1}{\mu} \int_{\underline{\sigma}}^{\bar{\sigma}} \{\beta_0 \log y_0(\sigma) + \beta_1 \log y_1(\sigma)\} d\sigma. \tag{4}$$

Recall that the subscripts 0 and 1 indicate innate ability ( $\tau = 0$ ) and training-intensive ( $\tau = 1$ ) tasks, respectively. We impose  $\sum_\tau \beta_\tau = 1$  and  $\mu \equiv \bar{\sigma} - \underline{\sigma}$  to ensure constant returns to scale.

Let there be a mass  $K$  of machine capital and normalize the labor force to have unit mass. We assume a skill distribution that is continuous and without mass points. Let  $V(s)$  denote the

---

<sup>12</sup>In reality, many of the innovations that lead to a fall in  $c_K$  may also cause a rise in  $A_K$ . However, the comparative statics with respect to  $A_K$  are qualitatively the same as those with respect to  $c_K$  (a proof is available upon request).

differentiable CDF, and  $v(s)$  the PDF, with support  $[\underline{s}, \bar{s}]$ . Factor market clearing conditions are

$$v(s) = \sum_{\tau} \int_{\underline{\sigma}}^{\bar{\sigma}} n_{\tau}(s, \sigma) d\sigma \quad \text{for all } s \in [\underline{s}, \bar{s}] \quad (5)$$

and

$$K = \sum_{\tau} \int_{\underline{\sigma}}^{\bar{\sigma}} k_{\tau}(\sigma) d\sigma. \quad (6)$$

Before characterizing the competitive equilibrium, we turn to a discussion of the assumptions underlying our task model.

## 2.5 Motivating the Assumptions of the Task Model

Researchers in artificial intelligence, robotics, and cognitive science have long been aware that some abilities that humans acquire quickly at an early age rely in fact on highly complex functions that are difficult if not impossible to reverse-engineer. In contrast, many abilities that humans must painstakingly acquire, such as mastery in arithmetic, are trivial from an engineering perspective. This observation has become known as Moravec’s paradox: “[It] is comparatively easy to make computers exhibit adult-level performance in solving problems on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility” (Moravec 1988, p.15). Moravec resolves the puzzle by considering the objective or intrinsic difficulty of a task, for instance the amount of information processing required, or the degrees of freedom and dexterity necessary to carry out a certain physical action. In terms of intrinsic difficulty, arithmetic is much easier than walking or face-to-face communication. A possible indicator of the intrinsic difficulty of a task, or its complexity, is the computer power a machine requires to perform the task. A common if imperfect measure of computer power is million instructions per second (MIPS).<sup>13</sup> UNIVAC I, a computer built in 1951 and able to perform arithmetic operations at a much faster rate than humans, performed at 0.002 MIPS. ASIMO, a robot introduced in 2000 that walks, recognizes faces and processes natural language, requires about 4,000 MIPS.<sup>14</sup> By this measure, a set of tasks routinely performed by any three-year-old is two million times more complex than arithmetic. The reason that we are usually not aware of this fact, and why Moravec’s observation at first seems puzzling, is that we rely on innate abilities<sup>15</sup> for functions like movement or perception, but have no such advantage when it comes to abstract tasks like arithmetic.<sup>16</sup>

<sup>13</sup>See Nordhaus (2007) for a discussion of MIPS as a measure of computer power.

<sup>14</sup>For a comparison of various computers (including UNIVAC I) and processors by MIPS, see [http://en.wikipedia.org/wiki/Instructions\\_per\\_second](http://en.wikipedia.org/wiki/Instructions_per_second), retrieved on October 16, 2013. For technical details of ASIMO, see Sakagami, Watanabe, Aoyama, Matsunaga, Higaki, and Fujimura (2002).

<sup>15</sup>“Innateness” of a certain skill does not need to imply that humans are born with it; instead, the subsequent development of the skill could be genetically encoded. For a critical discussion of the concept of innateness, see Mamelì and Bateson (2011).

<sup>16</sup>Moravec (1988, pp.15-16) provides an evolutionary explanation for this: “...survival in the fierce competition over such limited resources as space, food, or mates has often been awarded to the animal that could most quickly produce a correct action from inconclusive perceptions. Encoded in the large, highly evolved sensory and motor portions of the human brain is a billion years of experience about the nature of the world and how to survive

Table 1: Two-Dimensional Task Framework, Examples

	<i>Complexity</i>	
	–	+
<i>Little or no training required</i>	crushing rocks fast food preparation	customer reception waiting tables
<i>Training required</i>	bookkeeping weaving	lawyering management

We incorporate these insights into our framework by assuming that the amount of training a worker requires to be able to perform a task does not always depend on complexity.<sup>17</sup> Table 1 gives an overview of our task framework and contains examples. The bottom left and top right corners demonstrate the importance of a two-dimensional task space.

While we believe that our task framework is an improvement over existing literature, there are some limitations. For instance, technical change often leads to the introduction of new tasks and activities (flying airplanes, writing software). While our framework in principle allows for an endogenous task space, it does not suggest in what way technology might affect the set of tasks in the economy. Furthermore, automation does not necessarily involve machines replicating exactly the steps that humans carry out in completing a given task. Instead, a task can be made less complex by moving it to a more controlled environment.<sup>18</sup> Our framework does not explicitly allow for this possibility, but our conclusions should still be broadly correct if the cost of moving a process to a more controlled environment is increasing in its complexity. Finally, technological change tends to cause organizational change, but to keep the analysis tractable and to be able to focus on a single mechanism, we omit firm organization from the model.

What we do not view as a limitation is the assumption that machines could in principle perform any task. Comparative advantage ensures that some tasks will always be performed by humans, so that the model will be consistent with the fact that some tasks are not performed by machines in reality. More importantly, recent technological progress suggests that machine capabilities might be expanding quite rapidly. Brynjolfsson and McAfee (2011, p.14) argue that machines can potentially substitute for humans in a much larger range of tasks than was thought possible not long ago, citing recent advances in pattern recognition (driverless cars), complex communication (machine translation), and combinations of the two (IBM’s successful Jeopardy contestant Watson). Markoff (2012) provides an account of the increased flexibility, dexterity,

---

in it. The deliberate process we call reasoning is, I believe, the thinnest veneer of human thought, effective only because it is supported by this much older and much more powerful, though usually unconscious, sensorimotor knowledge. We are all prodigious olympians in perceptual and motor areas, so good that we make the difficult look easy. Abstract thought, though, is a new trick, perhaps less than 100 thousand years old. We have not yet mastered it. It is not all that intrinsically difficult; it just seems so when we do it.”

<sup>17</sup>Machines could be viewed as being endowed with some functions to the extent that materials have productive properties—take for instance copper with its electrical conductivity; but such endowments are usually highly specific and limited.

<sup>18</sup>See ALM (p.1283) and Simon (1960, pp.33-35). A recent example is the new sorting machine employed by the New York Public Library (Taylor 2010).

and sophistication of production robots. Shein (2013) gives examples of robots being increasingly adopted in manual tasks, such as collecting items in a warehouse, or the pruning of grapes.<sup>19</sup> For our model to be useful as a guide to medium-term future developments in the economy, we deem it prudent to make the most conservative assumption about what tasks are safe from automation.

## 2.6 Characterizing the Competitive Equilibrium

We normalize the price of the final good to one and denote the price of task  $(\sigma, \tau)$  by  $p_\tau(\sigma)$ . Profits of final good firms are given by

$$\Pi = Y - \sum_{\tau} \int_{\underline{\sigma}}^{\bar{\sigma}} p_\tau(\sigma) y_\tau(\sigma) d\sigma,$$

and profits of intermediate producers of task  $(\sigma, \tau)$  are

$$\Pi_\tau(\sigma) = p_\tau(\sigma) y_\tau(\sigma) - r k_\tau(\sigma) - \int_{\underline{s}}^{\bar{s}} w(s) n_\tau(s, \sigma) ds$$

where  $r$  is the rental rate of capital and  $w(s)$  is the wage paid to a worker with skill  $s$ . Design and training costs are included in intermediate producers' profits in the sense that for each unit of capital or labor hired, a fraction may be lost in design or training, as captured by (1), (2), and (3).

As in Costinot and Vogel (2010), a competitive equilibrium is defined as an assignment of factors to tasks such that all firms maximize profits and markets clear. Profit-maximizing task demand by final good producers is

$$y_\tau(\sigma) = \frac{\beta_\tau}{\mu} \frac{Y}{p_\tau(\sigma)}. \quad (7)$$

Profit maximization by intermediates producers implies

$$\begin{aligned} p_\tau(\sigma) \alpha^N(s, \sigma, \tau) &\leq w(s) \quad \forall s \in [\underline{s}, \bar{s}], \\ p_\tau(\sigma) \alpha^K(s_K, \sigma) &\leq r/A_K; \\ p_\tau(\sigma) \alpha^N(s, \sigma, \tau) &= w(s) \quad \text{if } n_\tau(s, \sigma) > 0, \\ p_\tau(\sigma) \alpha^K(s_K, \sigma) &= r/A_K \quad \text{if } k_\tau(\sigma) > 0. \end{aligned} \quad (8)$$

Denote the economy's set of tasks by  $T \equiv [\underline{\sigma}, \bar{\sigma}] \times \{0, 1\}$ . Formally, a *competitive equilibrium* in this economy is a set of functions  $y : T \rightarrow \mathbb{R}^+$  (task output);  $k : T \rightarrow \mathbb{R}^+$  and  $n : [\underline{s}, \bar{s}] \times T \rightarrow \mathbb{R}^+$  (factor assignment);  $p : T \rightarrow \mathbb{R}^+$  (task prices);  $w : [\underline{s}, \bar{s}] \rightarrow \mathbb{R}^+$  (wages); and a real number  $r$  (rental rate of capital) such that conditions (1) to (8) hold.

To be able to characterize the competitive equilibrium, we need to examine the properties of the productivity schedules  $\alpha^K$  and  $\alpha^N$ . In training-intensive tasks, workers face the same

---

<sup>19</sup>An overview of recent developments in basic robotics research can be found in Nourbakhsh (2013).

productivity schedule as machines, except for the skill parameter. Let  $\check{s} \in s_K \cup [s, \bar{s}]$  and define

$$\alpha(\check{s}, \sigma) \equiv 1 - \sigma/\check{s} \quad (\equiv \alpha^K(\check{s}, \sigma) \equiv \alpha^N(\check{s}, \sigma, \tau)). \quad (9)$$

Note that  $\alpha \in (0, 1)$ . Furthermore,  $\alpha_\sigma < 0$  and  $\alpha_{\check{s}} > 0$ . Productivity declines in complexity since a larger design or training expense is incurred. Higher skilled factors are more productive since they incur a smaller design or training expense. To characterize comparative advantage, we rely on the following result.

**Lemma 1** *The productivity schedule  $\alpha(\check{s}, \sigma)$  is strictly log-supermodular.*

The log-supermodularity of the productivity schedule implies that in training-intensive tasks, factors with higher skill have a comparative advantage in more complex tasks, or

$$\check{s}' > \check{s}, \sigma' > \sigma \quad \Leftrightarrow \quad \frac{\alpha(\check{s}', \sigma')}{\alpha(\check{s}, \sigma')} > \frac{\alpha(\check{s}', \sigma)}{\alpha(\check{s}, \sigma)}.$$

For instance, high skill workers have a comparative advantage over low skill workers in more complex tasks; all workers with  $s > s_K$  have a comparative advantage over machines in more complex tasks; and so on. The result is due to the fact that for higher skilled factors, training or design expenses increase less steeply with complexity.

Comparative advantage properties regarding training intensity are straightforward. Since  $\alpha$  is increasing in  $\check{s}$ , and because all workers have productivity one in all innate ability tasks, high skill workers have a comparative advantage over low skill workers in any training-intensive task. Furthermore, because machine productivity is the same in innate ability tasks as in training-intensive tasks if complexity is held constant, it follows that machines have a comparative advantage over all workers in any training-intensive task relative to the innate ability task with the same complexity. This seemingly trivial result has profound implications for the assignment of factors to tasks, and for the reallocation of factors in response to a fall in  $c_K$  (a rise in  $s_K$ ). It is at the root of the job polarization phenomenon, as we will show in Section 3 below.

The equilibrium assignment of factors to tasks is determined by comparative advantage, which is a consequence of the zero-profit condition (8).<sup>20</sup> Because high skill workers have a comparative advantage in training-intensive tasks (holding complexity constant), in equilibrium the labor force is divided into a group of low skill workers performing innate ability tasks, and a group of high skill workers carrying out training-intensive tasks: there exists a marginal worker with skill  $s^*$ , the least-skilled worker employed in training-intensive tasks. This is formally stated in part (a) of Lemma 2 below.

---

<sup>20</sup>To see how comparative advantage determines patterns of specialization, consider two firms, one producing training-intensive task  $\sigma$ , the other producing training-intensive task  $\sigma'$ . Suppose in equilibrium, firm  $\sigma$  is matched with workers of type  $s$  and firm  $\sigma'$  is matched with workers of type  $s'$ . Then (8) implies

$$\frac{\alpha(s', \sigma')}{\alpha(s, \sigma')} \geq \frac{\alpha(s', \sigma)}{\alpha(s, \sigma)},$$

which shows that type  $s$  ( $s'$ ) has a comparative advantage in task  $\sigma$  ( $\sigma'$ ), precisely the task to which she was assumed to be matched.

We focus on the empirically relevant case in which machines as well as workers perform both training-intensive and innate ability tasks.<sup>21</sup> In this case, machines are assigned to a subset of innate ability and training-intensive tasks that are relatively less complex, while low skill workers perform the remaining innate ability tasks: there is a threshold task  $\sigma_0^*$ , the marginal innate ability tasks, dividing the set of innate ability tasks into those performed by machines ( $\sigma \leq \sigma_0^*$ ) and those carried out by low skill workers ( $\sigma \geq \sigma_0^*$ ). Similarly, there is a marginal training-intensive task  $\sigma_1^*$  that divides the set of training-intensive tasks into those performed by machines ( $\sigma \leq \sigma_1^*$ ) and those carried out by high skill workers ( $\sigma \geq \sigma_1^*$ ). As in the case of the marginal worker, existence of these marginal tasks is of course a consequence of the comparative advantage properties discussed at the end of Section 2.3. These properties also imply  $\sigma_0^* < \sigma_1^*$ : the marginal training-intensive task is always more complex than the marginal innate ability task (recall that machines are relatively more productive in training-intensive tasks than workers, holding complexity constant); and  $s^* > s_K$ : it is always cheaper to train (though not to employ) the marginal worker than to design a machine in any task. These results are formally stated in part (b) of Lemma 2. An illustration of the equilibrium assignment is given in Figure 1.

**Lemma 2** (a) *In a competitive equilibrium, there exists an  $s^* \in (\underline{s}, \bar{s}]$  such that*

- $n_0(s, \sigma) > 0$  for some  $\sigma$  if and only if  $s \leq s^*$ , and
- $n_1(s, \sigma) > 0$  for some  $\sigma$  if and only if  $s \geq s^*$ .

(b) *If  $k_0(\sigma) > 0$  for some  $\sigma$ , then  $s^* > s_K$ , and there exist  $\sigma_0^*, \sigma_1^* \in [\underline{\sigma}, \bar{\sigma}]$  with  $\sigma_0^* < \sigma_1^*$  such that*

- $k_0(\sigma) > 0$  if and only if  $\sigma \leq \sigma_0^*$ ;
- $k_1(\sigma) > 0$  if and only if  $\sigma \leq \sigma_1^*$ ;
- $n_0(s, \sigma) > 0$  if and only if  $s \leq s^*$  and  $\sigma \geq \sigma_0^*$ ; and
- $n_1(s, \sigma) > 0$  if and only if  $s \geq s^*$  and  $\sigma \geq \sigma_1^*$ .

It remains to determine the assignment of low skill workers ( $s \leq s^*$ ) to innate ability tasks ( $\tau = 0, \sigma \geq \sigma_0^*$ ) and that of high skill workers ( $s \geq s^*$ ) to training-intensive tasks ( $\tau = 1, \sigma \geq \sigma_1^*$ ). The solution to the matching problem in innate ability tasks is indeterminate as all workers are equally productive in these tasks. However, knowledge of the assignment is not necessary to pin down task output and prices, as shown below. High skill workers are assigned to training-intensive tasks according to comparative advantage, with higher skilled workers carrying out more complex tasks. Formally, we have:

**Lemma 3** *In a competitive equilibrium, if  $s^* < \bar{s}$ , there exists a continuous and strictly increasing matching function  $M : [s^*, \bar{s}] \rightarrow [\sigma_1^*, \bar{\sigma}]$  such that  $n_1(s, \sigma) > 0$  if and only if  $M(s) = \sigma$ . Furthermore,  $M(s^*) = \sigma_1^*$  and  $M(\bar{s}) = \bar{\sigma}$ .*

<sup>21</sup>Sufficient conditions for the existence of such an equilibrium are derived Appendix B.1. We assume throughout that these conditions are satisfied. We note however that in general, it may happen that machines do not perform any innate ability tasks, and/or that workers do not carry out any training-intensive tasks.

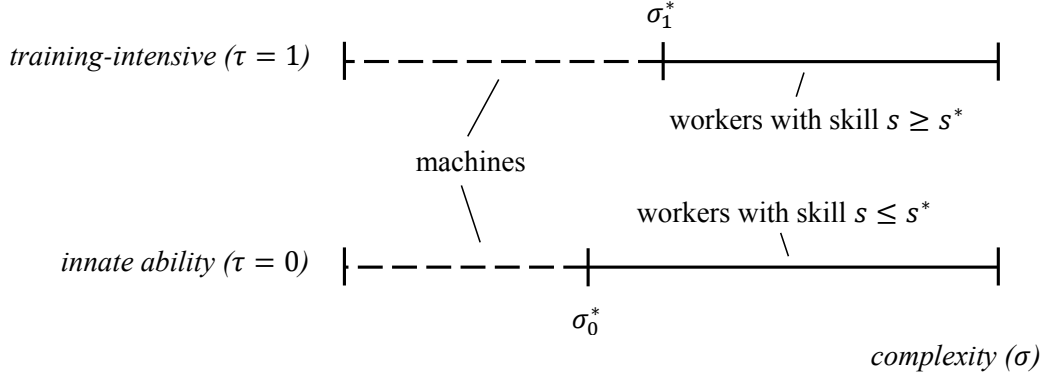


Figure 1: Assignment of labor and capital to tasks.

This result is an application of Costinot and Vogel (2010), with the added complication that domain and range of the matching function are determined by the endogenous variables  $s^*$  and  $\sigma_1^*$ . The matching function is characterized by a system of differential equations. Using arguments along the lines of the proof of Lemma 2 in Costinot and Vogel (2010), it can be shown that the matching function satisfies

$$M'(s) = \frac{\mu}{\beta_1} \frac{w(s)v(s)}{Y}, \quad (10)$$

and that the wage schedule is given by

$$\frac{d \log w(s)}{ds} = \frac{\partial \log \alpha(s, M(s))}{\partial s}. \quad (11)$$

The last equation is due to the fact that in equilibrium, a firm producing training-intensive task  $\sigma$  chooses worker skill  $s$  to minimize marginal cost  $w(s)/\alpha(s, \sigma)$ . Once differentiability of the matching function has been established, (10) can easily be derived from the market clearing condition (5) given Lemma 2, and using (7) and (8).<sup>22</sup> Figure 2 illustrates how the matching function assigns workers to training-intensive tasks.

In order to characterize the equilibrium more fully, and for comparative statics exercises, it is necessary to derive equations pinning down the endogenous variables  $\sigma_0^*$ ,  $\sigma_1^*$ , and  $s^*$ . These equations are due to a set of no-arbitrage conditions. In particular, firms producing the marginal

<sup>22</sup>In particular, Lemma 2 and (5) imply

$$\int_{s^*}^s v(s') ds' = \int_{\sigma_1^*}^{\sigma} n_1(M^{-1}(\sigma'), \sigma') d\sigma'.$$

Changing variables on the RHS of the last expression and differentiating with respect to  $s$  yields

$$v(s) = n_1(s, M(s))M'(s),$$

and substituting (3) we obtain

$$M'(s) = \frac{\alpha(s, M(s))v(s)}{y(M(s))}. \quad (12)$$

After eliminating task output and price using (7) and (8), (10) follows.

tasks are indifferent between hiring labor or capital, and the marginal worker is indifferent between performing innate ability tasks or the marginal training-intensive tasks. Formally, the price and wage functions must be continuous, otherwise the zero-profit condition (8) could not hold. This is a well-known result in the literature on comparative-advantage-based assignment models. Hence, the no-arbitrage conditions for the marginal tasks are

$$\frac{r}{A_K \alpha(s_K, \sigma_0^*)} = w(s) \quad \text{for all } s < s^* \quad (13)$$

and

$$\frac{r}{A_K \alpha(s_K, \sigma_1^*)} = \frac{w(s^*)}{\alpha(s^*, \sigma_1^*)}, \quad (14)$$

and the no-arbitrage condition for the marginal worker is

$$w(s) = w(s^*) \quad \text{for all } s \leq s^*. \quad (15)$$

The last result implies that there is a mass point at the lower end of the wage distribution. The mass point is a result of normalizing  $A$ , the amount of problems drawn, to one for all workers. To avoid the mass point, we could instead assume that  $A \equiv A(s)$  with  $A'(s) \geq 0$ . Equilibrium assignment and comparative statics results would be qualitatively the same. We maintain the normalization to avoid additional notation.

We can now complete the characterization of a competitive equilibrium by eliminating factor prices from (14). A standard implication of the Cobb-Douglas production function is that the mass of capital allocated to each task is constant within innate ability tasks and within training-intensive tasks. Some algebra shows<sup>23</sup> that machines produce task outputs

$$y_\tau(\sigma) = \frac{\beta_\tau A_K \alpha(s_K, \sigma) K}{\beta_0(\sigma_0^* - \underline{\sigma}) + \beta_1(\sigma_1^* - \underline{\sigma})} \quad \text{for all } \sigma \in [\underline{\sigma}, \sigma_\tau^*]. \quad (16)$$

Using these equations to solve for the task prices in (7), and plugging the obtained expression into (8), yields

$$r = \frac{\beta_0(\sigma_0^* - \underline{\sigma}) + \beta_1(\sigma_1^* - \underline{\sigma})}{\mu} \times \frac{Y}{K}. \quad (17)$$

---

<sup>23</sup>By (7) and (8), we have

$$\frac{y_\tau(\sigma)}{y_\tau(\sigma')} = \frac{\alpha(s_K, \sigma)}{\alpha(s_K, \sigma')}, \quad \frac{y_0(\tilde{\sigma})}{y_1(\tilde{\sigma}')} = \frac{\beta_0 \alpha(s_K, \tilde{\sigma})}{\beta_1 \alpha(s_K, \tilde{\sigma}')}$$

for any tasks  $(\sigma, \sigma', \tilde{\sigma}, \tilde{\sigma}')$  performed by machines. But (3), (3), and Lemma 2 imply

$$\frac{y_\tau(\sigma)}{y_\tau(\sigma')} = \frac{\alpha(s_K, \sigma) k_\tau(\sigma)}{\alpha(s_K, \sigma') k_\tau(\sigma')}, \quad \frac{y_0(\tilde{\sigma})}{y_1(\tilde{\sigma}')} = \frac{\alpha(s_K, \tilde{\sigma}) k_0(\tilde{\sigma})}{\alpha(s_K, \tilde{\sigma}') k_0(\tilde{\sigma}')}.$$

The previous two equations together give  $k_\tau(\sigma) = k_\tau(\sigma')$  and  $k_0(\tilde{\sigma}) = \frac{\beta_0}{\beta_1} k_1(\tilde{\sigma}')$ . By (6) and Lemma 2,

$$k_\tau(\sigma) = \frac{\beta_\tau K}{\beta_0(\sigma_0^* - \underline{\sigma}) + \beta_1(\sigma_1^* - \underline{\sigma})} \quad \text{for all } \sigma \in [\underline{\sigma}, \sigma_0^*].$$



This is of course the familiar result that with a Cobb-Douglas production function, factor prices equal the factor's share in output times total output per factor unit. In this case, the factor share is endogenously given by the (weighted) share of tasks to which the factor is assigned.

We employ similar steps to solve for  $w(s^*)$ . Since in innate ability tasks, worker productivity does not vary across tasks nor types, all innate ability tasks with  $\sigma \geq \sigma_0^*$  have the same price and all workers with  $s < s^*$  earn a constant wage equal to  $w(s^*)$  (as a result of the no-arbitrage condition for the marginal worker). As prices do not vary, neither does output, and so by the market clearing conditions (3) and (5),<sup>24</sup>

$$y_0(\sigma) = \frac{V(s^*)}{\bar{\sigma} - \sigma_0^*} \quad \text{for all } \sigma \geq \sigma_0^*. \quad (18)$$

Proceeding as above when solving for  $r$ , we obtain

$$w(s^*) = \frac{\beta_0(\bar{\sigma} - \sigma_0^*)}{\mu} \times \frac{Y}{V(s^*)}. \quad (19)$$

With (17) and (19) in hand, we can eliminate factor prices from the marginal cost equalization condition (13) to obtain

$$\frac{A_K \alpha(s_K, \sigma_0^*) K}{\beta_0(\sigma_0^* - \underline{\sigma}) + \beta_1(\sigma_1^* - \underline{\sigma})} = \frac{V(s^*)}{\beta_0(\bar{\sigma} - \sigma_0^*)}. \quad (20)$$

Also, combining conditions (13) to (15) yields

$$\alpha(s_K, \sigma_1^*) = \alpha(s_K, \sigma_0^*) \alpha(s^*, \sigma_1^*). \quad (21)$$

Lastly, (10) and (19) imply

$$M'(s^*) = \frac{\beta_0(\bar{\sigma} - \sigma_0^*)}{\beta_1} \frac{v(s^*)}{V(s^*)}. \quad (22)$$

Equations (4), (10), (11), (20), (21), and (22) together with the boundary conditions  $M(s^*) = \sigma_1^*$  and  $M(\bar{s}) = \bar{\sigma}$ , uniquely pin down the equilibrium objects  $\sigma_0^*$ ,  $\sigma_1^*$ ,  $s^*$ ,  $w$ , and  $M$ . The comparative statics analysis makes extensive use of these expressions.

To conclude this section, we highlight two properties of the wage structure in our model. First, integrating (11) yields an expression for the wage differential between any two skill types that are

---

<sup>24</sup>Under Lemma 2, integrating (5) yields

$$V(s^*) = \int_{\sigma_0^*}^{\bar{\sigma}} \int_{\underline{s}}^{s^*} n_0(s, \sigma) ds d\sigma,$$

but using (3) and the fact that task output is a constant  $y_0$  results in

$$V(s^*) = (\bar{\sigma} - \sigma_0^*) y_0.$$

both employed in training-intensive tasks,

$$\frac{w(s')}{w(s)} = \exp \left[ \int_s^{s'} \frac{\partial}{\partial z} \log \alpha(z, M(z)) dz \right] \quad \text{for all } s' \geq s \geq s^*. \quad (23)$$

This shows that wage inequality is fully characterized by the matching function (Sampson 2012). Second, adding (10) and (19) and integrating yields an expression for the average wage,

$$E[w] = \frac{\beta_0(\bar{\sigma} - \sigma_0^*) + \beta_1(\bar{\sigma} - \sigma_1^*)}{\mu} \times Y. \quad (24)$$

Since the labor force is normalized to have measure one, this expression also gives the total wage bill. It follows that the labor share in the model is given by the (weighted) share of tasks performed by workers.

### 3 Comparative Statics

Having outlined the model and characterized its equilibrium in the previous section, we now move on to comparative statics exercises. Our main interest is in investigating the effects of a fall in the machine design cost,  $c_K$ . In addition we will analyze the effects of increased skill abundance, motivated by the large increase in relative skill endowments seen in developed countries over the previous decades.

#### 3.1 Technical Change

Consider a fall in the machine design cost from  $c_K$  to  $\hat{c}_K$ , so that  $\hat{s}_K > s_K$ . Let  $M$  and  $\hat{M}$  be the corresponding matching functions, and similarly for  $\sigma_0^*$  and  $\hat{\sigma}_0^*$ ;  $\sigma_1^*$  and  $\hat{\sigma}_1^*$ ; and  $s^*$  and  $\hat{s}^*$ . We now state the main result of the paper.

**Proposition 1** *Suppose the machine design cost falls,  $\hat{c}_K < c_K$  and so  $\hat{s}_K > s_K$ . Then the marginal training-intensive task becomes more complex,  $\hat{\sigma}_1^* > \sigma_1^*$ , and the matching function shifts up,  $\hat{M}(s) > M(s)$  for all  $s \in [\max\{s^*, \hat{s}^*\}, \bar{s}]$ . If the fall in the machine design cost is such that  $\hat{s}_K \geq s^*$ , then the marginal worker becomes more skilled,  $\hat{s}^* > s^*$ .*

A fall in the machine design cost implies a rise in machine productivity and thus a fall in the marginal cost of employing machines in any task. Crucially, the marginal cost of employing machines in the threshold training-intensive tasks falls by more than the marginal cost in the threshold innate ability task, since  $\sigma_0^* < \sigma_1^*$ .<sup>25</sup> This means that machine employment in training-intensive tasks increases by more than in innate ability tasks. In fact, numerical simulations suggest that the effect of a fall in  $c_K$  on  $\sigma_0^*$  is ambiguous.

As machines are newly adopted in a subset of training-intensive tasks, the workers initially performing these tasks get replaced. Some of these workers upgrade to more complex training-intensive tasks—the matching function shifts up. Others downgrade to innate ability tasks, at

<sup>25</sup>Because  $\sigma_0^* < \sigma_1^*$  and due to the log-supermodularity of  $\alpha$ , the ratio  $\alpha(s_K, \sigma_1^*)/\alpha(s_K, \sigma_0^*)$  is increasing in  $s_K$ .

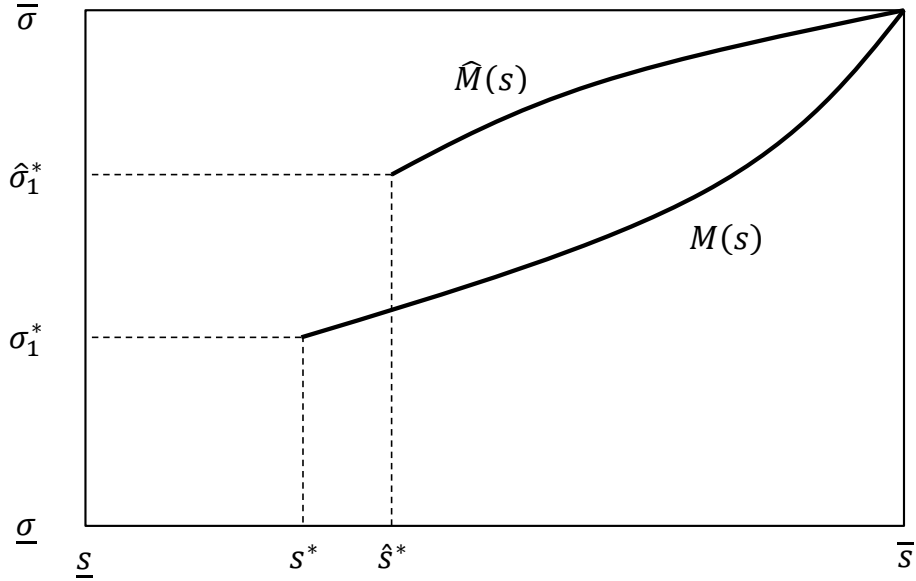


Figure 2: Assignment of workers to training-intensive tasks and the effects of technical change. Complexity  $\sigma$  is plotted on the vertical axis, while skill level  $s$  is plotted on the horizontal axis. The upward shift of the matching function and the shift of its lower end to the northeast are brought about by a fall in the machine cost from  $c_K$  to  $\hat{c}_K$  as stated in Proposition 1.

least for a sufficiently large fall in the machine design cost. Thus, labor-replacing technical change causes job polarization. These effects are illustrated by Figure 2.

While we are able to show that the threshold training-intensive task always becomes more complex and the matching function always shifts up, we are unable to rule out  $\hat{s}^* \leq s^*$  for small decreases in the machine design cost. However, we can prove that if machine design costs fall steadily over time, so that  $\hat{s}_K \geq s^*$  at some point, then the skill cutoff level must rise eventually. Thus, we limit our attention to the case in which a fall in  $c_K$  triggers a rise in  $s^*$  and hence job polarization occurs.

The matching function is a sufficient statistic for inequality (Sampson 2012), so that the shift in the matching function contains all the required information for deriving changes in relative wages among workers who remain in training-intensive tasks. Intuitively, as the upward shift implies skill downgrading by firms (but task upgrading for workers), the zero profit conditions imply that relatively low skill workers must have become relatively cheaper, or else their new employers would not be willing to absorb them. Hence the skill premium goes up for workers remaining in training-intensive tasks. Similar reasoning implies that workers who moved to innate ability tasks now earn relatively less than workers who were already performing these tasks. In sum, middle-skill workers who are displaced by machines experience downward pressure on their wages, and they end up worse off in relative terms compared to high and low skill workers who are less affected by automation (or not at all). Wage inequality rises at the top, but falls at the bottom of the distribution, as illustrated by Figure 3. The formal result is as follows.

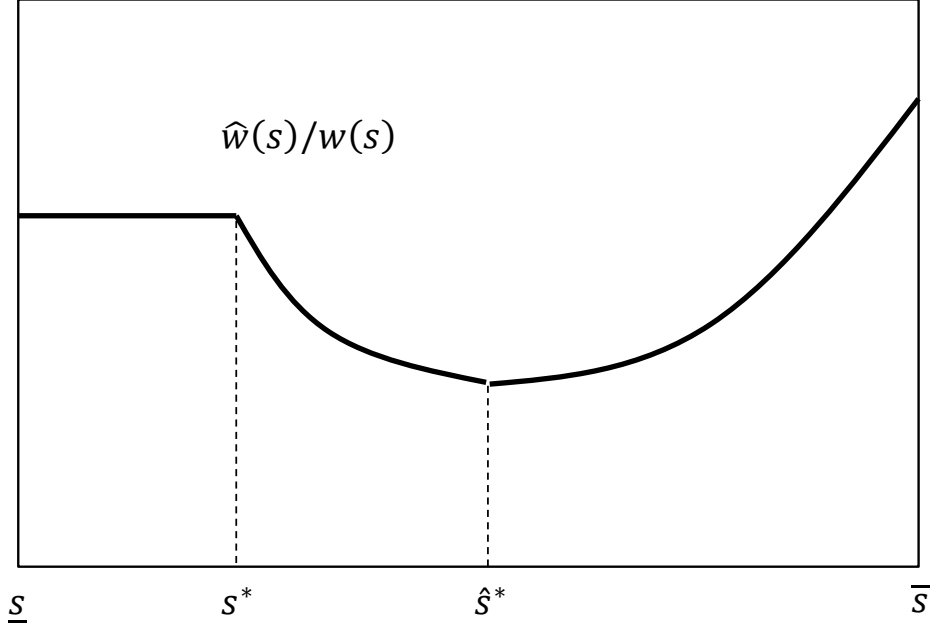


Figure 3: Changes in wages as a result of a fall in the machine design cost from  $c_K$  to  $\hat{c}_K$ . For each skill level  $s$ , the ratio of new to old wages is plotted. Workers with  $s \in [\hat{s}^*, \bar{s}]$  remain in training-intensive tasks and experience a rise in the skill premium. Workers with  $s \in [s^*, \hat{s}^*)$  switch to innate ability tasks and experience a fall in the skill premium. See Corollary 1 for details.

**Corollary 1** *Suppose  $\hat{c}_K < c_K$  and consider the case in which  $\hat{s}^* > s^*$ . Wage inequality increases at the top of the distribution but decreases at the bottom. Formally,*

$$\frac{\hat{w}(s')}{\hat{w}(s)} > \frac{w(s')}{w(s)} \quad \text{for all } s' > s \geq \hat{s}^*$$

and

$$\frac{\hat{w}(s')}{\hat{w}(s)} < \frac{w(s')}{w(s)} \quad \text{for all } s', s \text{ such that } \hat{s}^* > s' > s \geq s^*.$$

Relative wages are affected by technical change despite the fact that all factors are perfect substitutes at the task level. This is because tasks are  $q$ -complements in the production of the final good.<sup>26</sup> Intuitively, firms respond in two to the fall in the design cost. First, they upgrade existing machines. Second, they adopt machines in tasks previously performed by workers. The first effect on its own would lead to a rise in wages for all workers, because the increase in machines' task output raises the marginal product of all other tasks; moreover, relative wages would remain unchanged. The second effect, however, forces some workers to move to different tasks, putting downward pressure on their wages.<sup>27</sup> Since middle skill workers are most likely to be displaced by increased automation, their wages relative to low skill and high skill workers will

<sup>26</sup>This means that the price of a task increases in the output of all other task. The mechanism described in this paragraph has been highlighted by Acemoglu and Autor (2011).

<sup>27</sup>The effect works mainly through changes in task prices. Physical productivity actually increases for middle skill workers who get reassigned to innate ability tasks.

decline.<sup>28</sup> Thus, whether technology substitutes for or complements a worker of given skill type (in terms of relative wage effects) depends on that worker's exposure to automation, which is endogenous in our model.

To map the model's predictions for changes in wage inequality to the data, following Costinot and Vogel (2010) it is useful to distinguish between observable and unobservable skills. In particular, our continuous skill index  $s$  is unlikely to be observed by the econometrician. Instead, we assume that the labor force is partitioned according to some observable attribute  $e$ , which takes on a finite number of values and may index education or experience. Suppose further that high- $s$  workers are disproportionately found in high- $e$  groups. Formally, if  $s' > s$  and  $e' > e$ , we require  $v(s', e')v(s, e) \geq v(s, e')v(s', e)$ . Costinot and Vogel (2010) show that an increase in wage inequality in the sense of Corollary 1 implies an increase in the premium paid to high- $e$  workers as well as an increase in wage inequality among workers with the same  $e$ . In other words, the model predicts that if the machine design cost falls, both between and within (or residual) wage inequality will rise for the fraction of workers remaining in training-intensive tasks. On the other hand, within and between inequality falls for the set of workers below the new cutoff. This group includes stayers in as well as movers to innate ability tasks. In Section 5.2, we provide evidence consistent with this.

Although the effect on the marginal innate ability task is uncertain, the overall weighted share of tasks performed by machines increases. By (24), this is equivalent to a decrease in the labor share.

**Corollary 2** *Suppose  $\hat{c}_K < c_K$ . The labor share decreases.*

The fact that the labor share decreases means that it is not possible to sign the effect of a fall in the design cost on wage levels. Equation (24) shows that the average wage is affected both by the decrease in workers' task shares and the increase in output, so that the overall change is ambiguous. Of course, wage levels may also differentially change by worker type. For instance, high and low skill workers may enjoy absolute wage gains, while middle skill workers may suffer absolute wage losses.

### 3.2 Increase in Skill Abundance

Now consider an increase in the relative supply of skills. Following Costinot and Vogel (2010), we say that  $\hat{V}$  is more skill abundant relative to  $V$ , or  $\hat{V} \succeq V$ , if

$$\hat{v}(s')v(s) \geq \hat{v}(s)v(s') \quad \text{for all } s' > s.$$

Such a shift in the skill distribution implies first-order stochastic dominance and hence an increase in the mean. The shift may be due, for instance, to an increase in average education levels, to the extent that general education leads to attainment of knowledge applicable to a wide range of job tasks. If so, then training costs for a given level of complexity will fall on average, exactly as occurs if  $E[s]$  decreases.

---

<sup>28</sup>But middle skill workers' wages will not decline absolutely if the first effect dominates.

For simplicity, we restrict attention to distributions with common support, and we assume that  $\widehat{v}(\bar{s}) > v(\bar{s})$ . Characterizing comparative statics for changes in skill supplies is more challenging in our model than in the original Costinot-Vogel framework because domain and range of the matching function are endogenous. We are able to offer a partial result.

**Proposition 2** *Suppose that skill becomes more abundant,  $\widehat{V} \succeq V$  and  $\widehat{v}(\bar{s}) > v(\bar{s})$ . If this change in skill endowments induces an increase in the share of income accruing to labor, then the marginal training-intensive task becomes less complex,  $\widehat{\sigma}_1^* < \sigma_1^*$ ; the marginal worker becomes more skilled,  $\widehat{s}^* > s^*$ ; and the matching function shifts down,  $\widehat{M}(s) < M(s)$  for all  $s \in [\widehat{s}^*, \bar{s}]$ .*

Intuitively, such a change to the distribution of skills should raise the labor share, because the labor share equals the share of tasks performed by workers, and an increase in the average worker’s productivity should induce more firms to hire labor. While the labor share always increases in our numerical simulations, we are unable to prove the general result.<sup>29</sup>

The implications of Proposition 2 are as follows. Firms take advantage of the increased supply of skilled workers and engage in skill upgrading, which is equivalent to task downgrading for workers. This can be seen for training-intensive tasks by the downward shift of the matching function. For innate ability tasks, skill-upgrading is equivalent to the marginal worker becoming more skilled. Skill upgrading implies that the price of skill must have declined, so that the distribution of wages becomes more equal.

**Corollary 3** *Suppose  $\widehat{V} \succeq V$ , and that the labor share increases as a result. Then wage inequality decreases globally: for all  $s, s'$  with  $s' > s \geq s^*$ ,*

$$\frac{\widehat{w}(s)}{\widehat{w}(s')} > \frac{w(s)}{w(s')}.$$

The formal result implies a fall in both within and between inequality. This is consistent with a fall in the college premium induced by an increase in the supply of college educated workers as occurred in the US in the 1970s (Acemoglu 2002). Thus, our model features a modified version of the “Race between Education and Technology” (Goldin and Katz 2008), in the sense that education and technology have opposite effects on wage dispersion in the upper part of the distribution, but not in the lower part.

Proposition 2 says that the marginal training-intensive tasks becomes less complex, implying a decline in technology use for such tasks. In contrast, our simulations show that the marginal innate ability task becomes more complex. Thus, skill upgrading appears to coincide with technology being more (less) widely adopted in innate ability (training-intensive) tasks.

Equation (24) implies that changes in wages are in general not proportional to changes in GDP in this economy, unless task assignment is constant. By Corollary 2, technical change reduces the labor share, while numerical simulations suggest that an increase in skill supplies leads to an increase in the labor share. Thus, if both forces are at work, the labor share may not

<sup>29</sup>The labor share is given by  $\int_{\bar{s}}^{\widehat{s}} \frac{w(s)}{Y} dV(s)$ . Because  $\widehat{V}$  first-order stochastically dominates  $V$  and  $w(s)/Y$  is an increasing function, we have  $\int_{\bar{s}}^{\widehat{s}} \frac{w(s)}{Y} d\widehat{V}(s) > \int_{\bar{s}}^{\widehat{s}} \frac{w(s)}{Y} dV(s)$ . Thus, for the labor share to *decrease*, there would need to be a sufficiently large decline in wage-output ratios for a subset of workers.

exhibit any trend. We return to this issue when discussing existing empirical findings in Section 5.1.

## 4 Extensions

### 4.1 Making the Model Dynamic

Up to this point we have treated the economy's capital stock as exogenously given. To determine how endogenous capital accumulation would affect our comparative statics results, we assume that in the long run, the rental rate of capital is a constant pinned down by a time preference parameter<sup>30</sup> and that machines fully depreciate in every period. Furthermore, we assume that worker's knowledge depreciates fully in every period, or equivalently, there is an overlapping generations structure with each generation only working for one period. Suppose that the economy starts out in a steady state with the interest rate equal to its long-run value. Now recall that a fall in the machine design cost leads to a rise in the labor share. Furthermore, because the First Welfare Theorem applies to our model economy, output must not decrease, since the economy's resource constraint is less tight. By (17), we have that the interest rate increases. Thus, in the long run, the capital stock must increase to bring the interest rate back down.

It can be shown that a rise in the capital stock  $K$  has qualitatively the same effects on the marginal tasks, the matching function, and wages, as a fall in the machine design cost  $c_K$ .<sup>31</sup> This is because a higher supply of capital makes it cheaper to rent machines and thus encourages technology adoption. Thus, our predictions about the effects of a fall in  $c_K$  are not overturned with endogenous capital accumulation. In fact, the rise in the marginal training-intensive task, the upward shift of the matching function, the rise in the skill of the marginal worker, and the increase in wage inequality will be more pronounced in the long run as a result of the higher capital stock.

### 4.2 A Model with Fixed Costs

Our baseline model emphasizes that when a firm automates its production, total costs will generally be increasing in the firm's output and in the complexity of the processes required for production. While this in itself should be uncontroversial, our focus on variable costs with the implication of constant returns to scale is certainly restrictive. In particular, firms usually face large one-off expenses when installing new machinery.<sup>32</sup> While such expenses would generally depend on the scale at which the firm plans to operate, it is useful to consider the extreme case of a fixed setup cost.

In Appendix C we modify our baseline model such that firms wanting to automate production face a fixed cost (in units of the final good) which is increasing in the complexity of the task,

---

<sup>30</sup>Alternatively, we could assume that the economy is open to world capital markets, where it is a price taker.

<sup>31</sup>The proof is along similar lines as the proof of Proposition 1 and is available upon request. Since task-neutral machine productivity  $A_K$  enters the relevant model equations in the same way as  $K$ , the statement also applies to an increase in  $A_K$ .

<sup>32</sup>For an example relating to recent advances in AI, consider the concept of 'machine learning', where a software requires a considerable amount of initial 'training' before becoming operational.

but does not depend on the scale of production. We derive conditions ensuring an equilibrium assignment that is qualitatively the same as the one analyzed for the baseline model (see Figure 1). In particular, the marginal cost of using a machine must be sufficiently small, which can be achieved by making  $A_K$  very large, a realistic assumption; and the fixed cost must increase sufficiently in complexity. The model is much less tractable than the baseline model, and we are unable to derive general comparative statics results. Intuitively, when the fixed machine design cost falls, there is an incentive for firms to adopt machines in more-complex tasks. This incentive is stronger in training-intensive tasks: as complexity increases, the marginal cost of employing labor increases in training-intensive tasks but not in innate ability tasks. Thus, we would expect to see an increase in the share of workers performing innate-ability tasks. We have solved the model numerically and verified this intuition. We present results in Appendix C.

## 5 Empirical Support for the Model’s Predictions

Section 3.1 has established that any technological advance that facilitates automation of a wide range of tasks should lead to systematic shifts in task input, job polarization, and a hollowing out of the wage distribution. In addition, the model also predicts which worker types will be replaced as more tasks are automated, and to which task a displaced worker gets reassigned. In this section we briefly review papers that document these patterns for the recent information and communication technology revolution. We then discuss two studies presenting historical evidence that we also find to be consistent with the model’s prediction. Finally, we present new evidence consistent with our model’s predictions about trends in worker training levels.

### 5.1 Existing Evidence

*Changes in task input.*—In a seminal contribution, Autor, Levy, and Murnane (2003) document a decline in the fraction of workers performing “routine tasks”, and show that this decline is larger in industries that more rapidly adopted information technologies. They also find that “non-routine” interactive and analytic task inputs increased, and more so in industries with more rapid ICT adoption. Although routine-ness is conceptually distinct from complexity, ALM’s empirical measures of routine-ness may in fact be correlated with it. For example, they classify routine occupations as those that require “finger dexterity” and “adaptability to situations requiring the precise attainment of set limits, tolerances or standards.” It is likely that these are occupations with low complexity (though not necessarily low training intensity). The measured shift away from routine tasks is then consistent with our prediction of a reallocation towards more-complex tasks.

*Job polarization.*—Goos and Manning (2007) were the first to suggest that the “de-routinization” documented by ALM implies a polarization of employment since routine tasks were traditionally performed by middle-skill workers. They do find evidence of job polarization for the UK, and subsequently Autor, Katz, and Kearney (2006) showed this to be the case in the US as well. Goos, Manning, and Salomons (2009) provide evidence for job polarization in a majority of European economies, and show that much of it can be attributed to tasks shifts consistent with technical



change being the driving force. Importantly, Michaels, Natraj, and Van Reenen (forthcoming) show that in a sample of several developed countries it is indeed the case that industries that invested more heavily in information and communication technologies witnessed a decline in relative middle skill employment and wage bills, confirming the link between technical change and job polarization.

Cortes (2012) uses panel data from the US and shows that worker ability is a strong determinant of the destination occupation for workers exiting from routine occupations. He shows that low (high) ability workers are more likely to switch to non-routine manual (non-routine) cognitive occupations. This is consistent with our model if we interpret non-routine manual as innate ability tasks and non-routine cognitive as training-intensive and complex tasks.

*Wages.*—Recall that Corollary 1 implies a fall in wage inequality at the bottom of the distribution and a rise at the top. Consistent with this, Autor, Katz, and Kearney (2006) document that in the US since the late 1980s, wages in the middle of the distribution have risen more slowly than those at the top and bottom. Dickens, Manning, and Butcher (2012) show similar evidence for the UK and argue that the compression of the lower part of the distribution is partly explained by rises in the minimum wage. We interpret this as leaving room for a technological explanation like the one proposed here.

Lemieux (2006) shows that in the 1990s increases in within-group inequality were concentrated in the upper part of the wage distribution. For between-group wage differentials, Lindley and Machin (2011) document that in addition to a rise in the college premium, there has also been an increase in the wages of workers with a graduate degree relative to those with college only. Similarly, Angrist, Chernozhukov, and Fernández-Val (2006) document a more pronounced rise in within-group inequality for college graduates than for high school graduates, and an increase in the effect of an additional year of schooling on the upper tail of the conditional wage distribution, relative to the effect on lower tail and median. Thus, the evidence on within- and between-group inequality appears consistent with our model.

Firpo, Fortin, and Lemieux (2011) investigate using US data whether changes in the wage distribution can be attributed to changes in the returns to tasks that are due to technical change or offshoring. They find a prominent role of technology, while offshoring has become more important in the most recent decade. However, their identification assumptions may be viewed as restrictive from the perspective of our model, so that further research is required. Cortes (2012), in addition to providing evidence on worker movements, also shows that relative wages of those workers staying in middle-wage, routine occupations decline. Boehm (2013) uses NLSY data to estimate workers' selection into occupations based on observed comparative advantage. He finds that workers with a comparative advantage in routine occupations saw their wages decline relative to other workers, and even absolutely. Overall, the evidence on wages appears consistent with our model.

*Labor share.*—Recall that our comparative statics suggest the absence of a trend in the labor share when technical change and increases in skill supply occur simultaneously. However when technical change dominates, the labor share should fall. Rodriguez and Jayadev (2010) document that the labor share started to decline in most countries around 1980. Karabarbounis and Neiman

(2013) find that a substantial part of the decline can be explained by advances in ICT, which may be consistent with our model, although the structural model they estimate is quite different from ours. Elsby, Hobijn, and Sahin (2013) do not find a strong role of technology in explaining the decline in the labor share. Thus, there appears to be no consensus on this question yet, and further research is required.

*Historical evidence.*—Katz and Margo (2013) find that from 1850 to 1880, US manufacturing witnessed a relative decline in middle skill jobs like artisans at the expense of high skill jobs (non-production workers) and low skill jobs (operatives), concurrent with the increased adoption of steam power. Gray (2013) shows that electrification in the US during the first half of the 20th century led to a fall in the demand for dexterity-intensive tasks performed by middle skill workers, relative to manual and clerical tasks performed by low and high skill workers, respectively. These papers provide support for the model’s implication that job polarization is not a unique consequence of the ICT revolution.

Bessen (2011) provides evidence on weavers employed at a 19th century Massachusetts firm that gradually increased the degree of mechanization during the period studied. Even though some of workers’ skills were no longer needed as more tasks were automated, the tasks to which workers were reassigned required substantial on-the-job learning, much like the reassignment of workers to more-complex, training-intensive tasks in our model. Crucially, worker productivity in the remaining tasks increased, supporting the assumption of  $q$ -complementarity of tasks that underlies our model. Note that we would not necessarily expect an aggregate phenomenon like job polarization to occur at the firm level.

## 5.2 Trends in Occupational Training Requirements

In the model, training levels vary systematically with task characteristics. In particular, more complex training-intensive tasks require more training. We view occupations as bundles of tasks, so that a given occupation may combine tasks from across the task space. Measures of occupational characteristics should be informative about which region of the task space features most prominently in a given occupation. Thus, occupations with low training requirements should be intensive in innate ability tasks; and occupations with very high training requirements should feature highly complex, training-intensive tasks.

To measure training requirements of occupations, we use the Fourth Edition Dictionary of Occupational Titles (DOT) in combination with the 1971 April Current Population Survey (CPS) (National Academy of Sciences 1981), and the US Department of Labor’s O\*NET database in combination with the 2008 American Community Survey (ACS). The information in the 2008 ACS refers to the previous year. Hence, our data cover the years 1971 and 2007. Since the 1971 April CPS lacks information on earnings, we also used the IPUMS 1970 census extract which contains earnings data pertaining to 1969.<sup>33</sup> We use David Dorn’s three-digit occupation codes throughout (Dorn 2009). Our analysis is based on a sample of all employed persons aged 17 to 65. To see whether our results are driven by changes in composition, we repeated the analysis using a

---

<sup>33</sup>Because we have to merge separate data sets at the three-digit occupation level, we prefer using the census to the much smaller 1971 March CPS for obtaining earnings data.

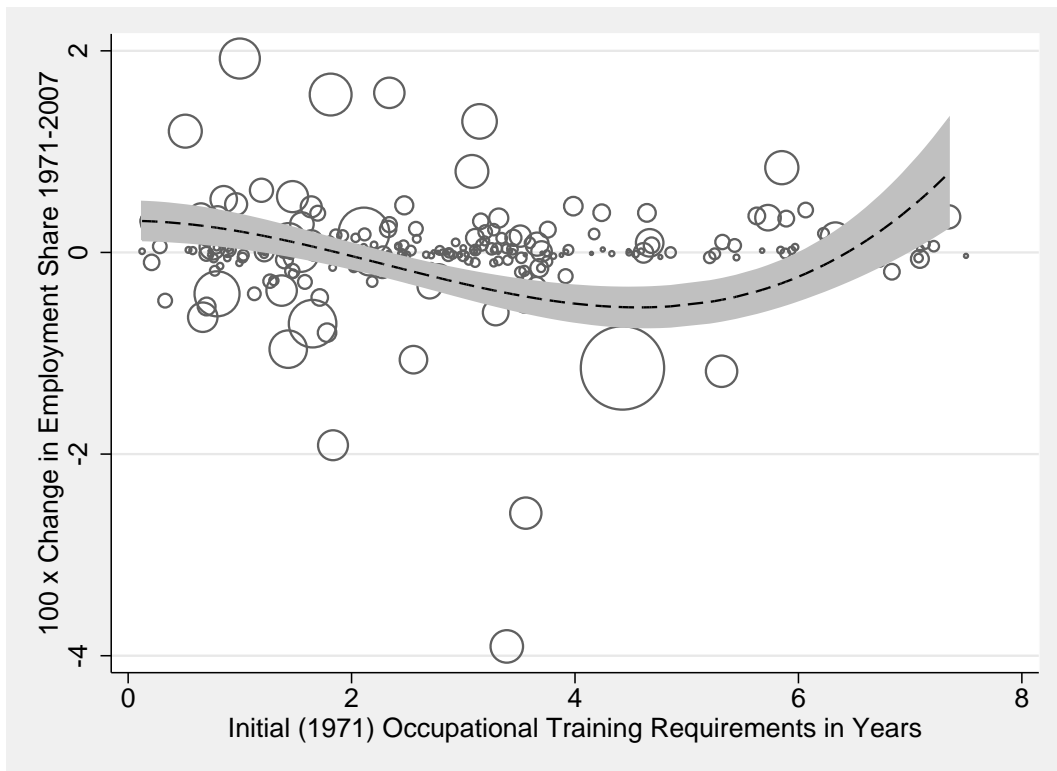


Figure 4: Changes in occupational shares by initial training requirements. We calculate training requirements using the variable *specific vocational preparation* (SVP) from the Dictionary of Occupational Titles and the O\*NET database. Observations are weighted by average occupational employment shares. Fitted curves are fractional polynomials, drawn using Stata's *fpfitci* option.

sample of white males only. The results, available upon request, are qualitatively identical.

Both the DOT and O\*NET contain the variable *Specific Vocational Preparation* (SVP), which indicates “the amount of time required to learn the techniques, acquire the information, and develop the facility needed for average performance in a specific job-worker situation. SVP includes training acquired in a school, work, military, institutional, or vocational environment, but excludes schooling without specific vocational content” (National Academy of Sciences 1981, p.21 in codebook). SVP is a bracketed variable and we use midpoints to convert it into training time measured in years. See Appendix D for details. Tables D.3 and D.4 list the twenty most and least training-intensive occupations in 1971 and 2007, respectively.

The definition of SVP matches our concept of task-specific training more closely than years of education. This is because much of education, at least up to high school graduation, is general in nature and the skills acquired are portable across occupations. Also, the average level of education of workers in a given occupation may be affected by the supply of educated workers independently of actual training requirements—we provide evidence for this below. In professional occupations such as lawyers and physicians there is a clear mapping between years of schooling and training requirements, but in general this is not the case. In terms of our model, we think of general education as affecting the ability to acquire task-specific knowledge. Thus, years of schooling may proxy for  $s$ .

The model delivers several predictions about trends in training requirements. First, as a fall in the machine design cost triggers a reallocation of workers towards tasks of higher complexity on the one hand (the upward shift of the matching function) and towards innate ability tasks on the other, the model predicts a polarization of job training requirements. Figure 4 plots fitted values from a locally weighted regression of changes in an occupation’s employment share on its percentile rank in the 1971 distribution of occupational mean wages.<sup>34</sup> The pattern is consistent with the model’s prediction of polarization of training requirements.

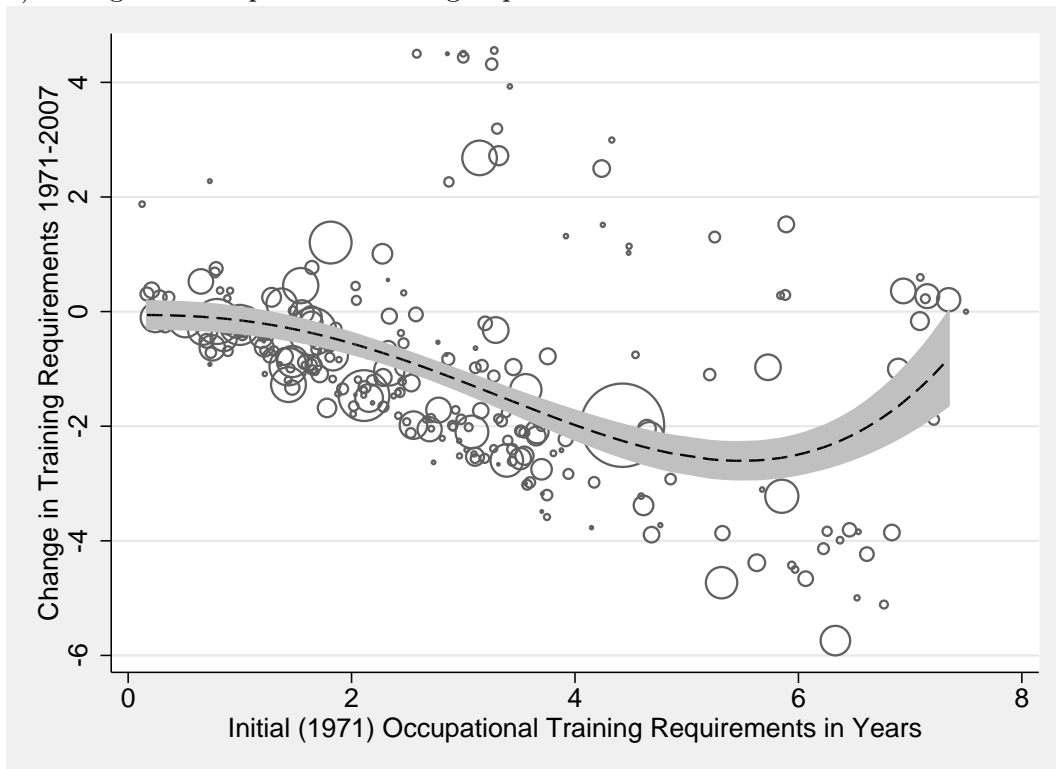
Second, the model can potentially help to make sense of changes in training requirements within occupations. Let us assume that measured training requirements are indicative of the most complex training-intensive task within an occupation’s bundle of tasks. Then we would expect training requirements to decrease in occupations with intermediate training requirements, as the most complex tasks in these occupations are automated. Panel a) of Figure 5 shows that indeed, occupations with intermediate initial training requirements saw the largest declines in training requirements. These occupations include air traffic controllers, precision makers, insurance adjusters, and various engineering occupations (see Table D.5), which appears consistent with our automation-based explanation.

Third, our model predicts that an increase in the supply of general skill  $s$  should result in skill upgrading across tasks. Indeed, average years of schooling increased in almost all occupations, as shown in panel b) of Figure 5. Furthermore, changes in occupation average years of schooling do not follow the same pattern as changes in training requirements, supporting our assertion that the two measures relate to distinct concepts.

---

<sup>34</sup>We employ the same estimation method as Acemoglu and Autor (2011) and Autor and Dorn (2013) to facilitate comparison with their plots of employment share changes against initial occupational mean wages.

a) Changes in occupational training requirements



b) Changes in occupational average years of schooling



Figure 5: Changes in occupational training requirements and average years of schooling. Observations are weighted by average occupational employment shares. Fitted curves are fractional polynomials, drawn using Stata's *fpfitci* option.



Figure 6: Growth of occupational labor input against changes in training requirements. Fitted line from a regression of changes in log total hours on changes in log training requirements. The estimated coefficient is 0.33 with a robust standard error of 0.08. Observations are weighted by average occupational employment shares.



Figure 7: Changes in occupational mean wages against changes in training requirements. Occupational mean wages have been adjusted for sex, race, experience, and region. Fitted line from a regression of changes in mean wages on changes in log training requirements. The estimated coefficient is 0.070 with a robust standard error of 0.026. Observations are weighted by average occupational employment shares.

Fourth, if decreases in training requirements are due to increased automation, then employment growth should have been lower in occupations with larger decreases in training requirements. This is indeed the case. A regression of changes in log total hours on changes in log training requirements yields a coefficient of 0.33 (robust standard error 0.08). Raw data and fitted line are plotted in Figure 6. Including changes in log years of education on the right hand side slightly decreases the coefficient on training.<sup>35</sup>

Fifth, we consider how changes in training requirements correlate with changes in occupational mean wages. We obtain adjusted occupational mean log wages as the predicted values from a regression of log wages on occupation dummies, a quartic in potential experience, region dummies, and indicators for female and non-white, evaluated at sample means. A regression of changes in occupation log wages on changes in log training requirements yields a coefficient of 0.07 (standard error 0.026), see Figure 7. Including changes in log years of education on the right hand side slightly increases the coefficient on training.

The finding is consistent with the model if we interpret falls in training requirements as

<sup>35</sup>A positive and statistically significant relationship also exists between employment growth and changes in the level of training requirements; and between changes in occupational employment shares and changes in both the level and log of training requirements.

Table 2: Regressions of Log Weekly Wages on Years of Schooling By Training Requirements

	(1)	(2)	(3)	(4)
	Training $\leq$ 6 months		Training $>$ 6 months	
	1969	2007	1969	2007
Full sample	0.053 (0.003)	0.031 (0.001)	0.076 (0.000)	0.115 (0.000)
White males	0.060 (0.003)	0.036 (0.001)	0.074 (0.000)	0.110 (0.000)

Note: Coefficients on years of schooling are shown. Data for columns (1)-(2) include workers in occupations requiring six months of training or less, data for columns (3)-(4) include the remaining workers. Training data are from the 4th edition of the Dictionary of Occupational Titles and the O\*NET database. Wage data are from the 1970 census and the 2008 American Community Survey, respectively. Robust standard errors in parentheses. Regressions are weighted by the relevant sampling weights. Controls include a quartic in experience, dummies for female and race other than white, and four regional dummies.

increased automation of tasks. For concreteness, consider an occupation whose task bundle initially includes training-intensive tasks with complexities between  $\sigma_1^*$  and  $\sigma' > \widehat{\sigma}_1^*$ . Let  $s'$  be the skill level of the worker initially performing task  $\sigma'$ . After the fall in machine design costs, all tasks in the interval  $[\sigma_1^*, \widehat{\sigma}_1^*]$  are newly automated. Workers with skill levels between  $\widehat{s}^*$  and some  $s'' < s'$  will remain in the occupation. Figure 3 shows that these workers experience wage declines relative to most other workers.

Finally, we test the model’s prediction about changes between-group inequality. In particular, we consider changes in returns to schooling. As discussed in Section 3, a fall in the design cost should lead to a rise in returns for workers above the new cutoff  $\widehat{s}^*$ , and to a fall in returns for workers below the cutoff. We implement this by running Mincer regressions separately for workers in occupations requiring six months of training or less and for all remaining workers. To address concerns related to changes in composition, we also show results for white males only.

The evidence on returns to schooling is consistent with the model’s predictions. As Table 2 shows, returns fell for workers in occupations requiring little training, but rose for the remaining workers. Note that for these results to constitute a valid test of the model, we do not need to rely on a causal interpretation of the Mincer regressions. The model makes a prediction about how conditional means of the wage distribution should change as a result of differential changes to labor demand and worker movements. We employ regressions to describe these conditional means.

## 6 Conclusion

In this paper we explain labor market polarization as resulting from technological advances that allow firms to more cheaply automate tasks. Our model thus explains why polarization occurred not only following the information and communication technology revolution, but also after the



introduction of the steam engine in the US in the 19th century and during the electrification of US manufacturing in the early 20th century. The model explains why firms do not automate processes although it would be feasible to do so, as in the case of fast food preparation, and why many arguably more complex processes, such as in manufacturing, have already been automated. The key is that difficulty in an engineering sense is not necessarily correlated with difficulty in terms of the amount of training that a human requires to complete a task. Our model delivers novel predictions about changes in training requirements, which we find to be consistent with US data. We have not allowed for changes in the economy's task mix or changes in firm organization resulting from technical change—further research is necessary to determine whether our results are robust to these extensions.

## References

- ACEMOGLU, D. (2002): “Technical Change, Inequality, and the Labor Market,” *Journal of Economic Literature*, 40(1), 7–72.
- ACEMOGLU, D., AND D. AUTOR (2011): “Skills, Tasks and Technologies: Implications for Employment and Earnings,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4, chap. 12, pp. 1043–1171. Elsevier.
- ANGRIST, J., V. CHERNOZHUKOV, AND I. FERNÁNDEZ-VAL (2006): “Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure,” *Econometrica*, 74(2), 539–563.
- ATAK, J., F. BATEMAN, AND T. WEISS (1980): “The Regional Diffusion and Adoption of the Steam Engine in American Manufacturing,” *The Journal of Economic History*, 40(02), 281–308.
- AUTOR, D. H. (2013): “The task to labor markets: an overview,” *Journal for Labour Market Research*, 46(3), 185–199.
- AUTOR, D. H., AND D. DORN (2013): “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market,” *American Economic Review*, 103(5), 1553–97.
- AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2006): “The Polarization of the U.S. Labor Market,” *American Economic Review*, 96(2), 189–194.
- AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): “The Skill Content Of Recent Technological Change: An Empirical Exploration,” *The Quarterly Journal of Economics*, 118(4), 1279–1333.
- BESSEN, J. (2011): “Was Mechanization De-Skilling? The Origins of Task-Biased Technical Change,” Working Papers 1101, Research on Innovation.
- BOEHM, M. J. (2013): “The Wage Effects of Job Polarization: Evidence from the Allocation of Talents,” mimeo, LSE.
- BOFF, R. B. D. (1967): “The Introduction of Electric Power in American Manufacturing,” *The Economic History Review*, 20(3), pp. 509–518.
- BRYNJOLFSSON, E., AND A. MCAFEE (2011): *Race Against The Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*. Digital Frontier Press.
- CASELLI, F. (1999): “Technological Revolutions,” *American Economic Review*, 89(1), 78–102.
- CHARI, V. V., AND H. HOPENHAYN (1991): “Vintage Human Capital, Growth, and the Diffusion of New Technology,” *Journal of Political Economy*, 99(6), 1142–65.
- CORTES, G. M. (2012): “Where Have the Routine Workers Gone? A Study of Polarization Using Panel Data,” The School of Economics Discussion Paper Series 1224, Economics, The University of Manchester.
- COSTINOT, A. (2009): “On the origins of comparative advantage,” *Journal of International Economics*, 77(2), 255–264.
- COSTINOT, A., AND J. VOGEL (2010): “Matching and Inequality in the World Economy,” *Journal of Political Economy*, 118(4), 747–786.
- DAVIDSON, A. (2012): “Making It in America,” *The Atlantic*, Jan/Feb.
- DICKENS, R., A. MANNING, AND T. BUTCHER (2012): “Minimum Wages and Wage Inequality: Some Theory and an Application to the UK,” Working Paper Series 4512, Department of Economics, University of Sussex.
- DORN, D. (2009): “Essays on Inequality, Spatial Interaction, and the Demand for Skills,” Dissertation no. 3613, University of St. Gallen.
- DORNBUSCH, R., S. FISCHER, AND P. A. SAMUELSON (1977): “Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods,” *American Economic Review*, 67(5), 823–39.
- ELSBY, M. W. L., B. HOBIJN, AND A. SAHIN (2013): “The Decline of the U.S. Labor Share,” Fall 2013 Brookings Panel on Economic Activity.

- FIRPO, S., N. M. FORTIN, AND T. LEMIEUX (2011): "Occupational Tasks and Changes in the Wage Structure," IZA Discussion Papers 5542, Institute for the Study of Labor (IZA).
- GARICANO, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108(5), 874–904.
- GARICANO, L., AND E. ROSSI-HANSBERG (2006): "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 121(4), 1383–1435.
- GOLDIN, C., AND L. F. KATZ (1998): "The Origins Of Technology-Skill Complementarity," *The Quarterly Journal of Economics*, 113(3), 693–732.
- (2008): *The Race between Education and Technology*. Harvard University Press.
- GOOS, M., AND A. MANNING (2007): "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *The Review of Economics and Statistics*, 89(1), 118–133.
- GOOS, M., A. MANNING, AND A. SALOMONS (2009): "Job Polarization in Europe," *American Economic Review*, 99(2), 58–63.
- (2011): "Explaining job polarization: the roles of technology, offshoring and institutions," Open Access publications from Katholieke Universiteit Leuven urn:hdl:123456789/331184, Katholieke Universiteit Leuven.
- GRAY, R. (2013): "Taking technology to task: The skill content of technological change in early twentieth century United States," *Explorations in Economic History*, 50(3), 351–367.
- HOLMES, T. J., AND M. F. MITCHELL (2008): "A theory of factor allocation and plant size," *RAND Journal of Economics*, 39(2), 329–351.
- JOVANOVIC, B. (1998): "Vintage Capital and Inequality," *Review of Economic Dynamics*, 1(2), 497–530.
- KARABARBOUNIS, L., AND B. NEIMAN (2013): "The Global Decline of the Labor Share," NBER Working Papers 19136, National Bureau of Economic Research, Inc.
- KATZ, L. F., AND R. A. MARGO (2013): "Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective," NBER Working Papers 18752, National Bureau of Economic Research, Inc.
- LEMIEUX, T. (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, 96(3), 461–498.
- LEVY, F., AND R. J. MURNANE (2004): *The New Division of Labor: How Computers Are Creating the Next Job Market*. Princeton University Press.
- LINDLEY, J., AND S. MACHIN (2011): "Rising Wage Inequality and Postgraduate Education," IZA Discussion Papers 5981, Institute for the Study of Labor (IZA).
- MAMELI, M., AND P. BATESON (2011): "An evaluation of the concept of innateness," *Phil. Trans. R. Soc. B*, 366, pp. 436–443.
- MARKOFF, J. (2012): "Skilled Work, Without the Worker," *The New York Times*, Aug 18.
- MELENDEZ, E. D. (2013): "Minimum Wage Jobs Won't Be Replaced With This Robot Anytime Soon," *The Huffington Post*, Aug 30.
- MICHAELS, G., A. NATRAJ, AND J. VAN REENEN (forthcoming): "Has ICT Polarized Skill Demand? Evidence from Eleven Countries over 25 Years," *Review of Economics and Statistics*, (16138).
- MORAVEC, H. (1988): *Mind Children*. Harvard University Press.
- NATIONAL ACADEMY OF SCIENCES (1981): "Dictionary of Occupational Titles (DOT): Part I - Current Population Survey, April 1971, Augmented With DOT Characteristics and Dictionary of Occupational Titles (DOT): Part II - Fourth Edition Dictionary of DOT Scores for 1970 Census Categories," .
- NORDHAUS, W. D. (2007): "Two Centuries of Productivity Growth in Computing," *The Journal of Economic History*, 67(01), 128–159.
- NOURBAKSH, I. R. (2013): *Robot Futures*. MIT Press.

- RODRIGUEZ, F., AND A. JAYADEV (2010): “The Declining Labor Share of Income,” Human Development Research Papers (2009 to present) HDRP-2010-36, Human Development Report Office (HDRO), United Nations Development Programme (UNDP).
- SAKAGAMI, Y., R. WATANABE, C. AOYAMA, S. MATSUNAGA, N. HIGAKI, AND K. FUJIMURA (2002): “The intelligent ASIMO: System overview and integration,” Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems.
- SAMPSON, T. (2012): “Selection into Trade and Wage Inequality,” CEP Discussion Papers dp1152, Centre for Economic Performance, LSE.
- SHEIN, E. (2013): “Businesses adopting robots for new tasks,” *Computerworld*, Aug 1.
- SIMON, H. (1960): “The corporation: Will it be managed by machines?,” in *Management and corporations*, ed. by M. Anshen, and G. Bach, pp. 17–55. McGraw-Hill.
- TAYLOR, K. (2010): “That Mighty Sorting Machine Is Certainly One for the Books,” *The New York Times*, Apr 22.
- ZEIRA, J. (1998): “Workers, Machines, And Economic Growth,” *The Quarterly Journal of Economics*, 113(4), 1091–1117.

# Appendices

## A An Extended Model of Task Production and Firms' Productivity Choices

Here we model the production process for tasks explicitly, following Garicano (2000). In order to produce, factors (workers, machines) must confront and solve problems. These problems are task-specific. There is a continuum of problems  $Z \in [0, \infty)$  in each task, and problems are ordered by frequency. Thus, there exists a non-increasing probability density function for problems in each task.

Factors draw problems and produce if and only if they know the solution to the problem drawn. We assume that a mass  $A$  of problems is drawn, and  $A$  may vary across factors. Hence, the task-neutral productivity term introduced in Section 2.2 has a more precise interpretation in this context. Task output per factor unit is equal to  $A$  times the integral of the density function over the set of problems to which the factor knows the solution.

For simplicity, we will again assume that all workers draw a unit mass of problems in all tasks, or  $A = 1$ . Equilibrium assignment and comparative statics results are qualitatively the same if we instead assume that  $A \equiv A(s)$  with  $A'(s) \geq 0$ .

The distribution of problems in a task with complexity  $\sigma$  is given by the cumulative distribution function  $F(Z; \sigma)$ , which we assume to be continuously differentiable in both  $Z$  and the shift parameter  $\sigma$ . Let  $\partial F / \partial \sigma < 0$ , so that  $\sigma$  indexes first-order stochastic dominance. In terms of the some of the examples discussed in Section 2.5, driving a car and grading an essay are more complex (higher  $\sigma$ ) than driving a train or grading an MCQ test since the number of distinct problems typically encountered in the former set of tasks is higher than in the latter.

The probability density function corresponding to  $F$  is  $f(Z; \sigma)$ . Because  $F$  is continuously differentiable and  $Z$  indexes frequency,  $f$  is strictly decreasing in  $Z$ . We impose the following condition on the family of distributions  $F(Z, \sigma)$ .

**Assumption 1**  $F(z, \sigma)$  is strictly log-supermodular.

This assumption will give rise to the same comparative advantage properties as in the baseline model. One of the distributions satisfying Assumption 1 is the exponential distribution with mean  $\sigma$ .

Note that the distribution of problems depends only on  $\sigma$  and not on  $\tau$ . As discussed above, training intensity is not an intrinsic property of a task, but arises from the fact that humans have evolved such that some tasks require less effort to master than others, even holding constant (objective) complexity. In this context, humans are assumed to be endowed with knowledge of the solutions to all problems in innate ability tasks.

We now characterize optimal training and design choices and derive equilibrium productivity of workers and machines. First observe that firms will equip factors with a set of knowledge  $[0, z]$ , since it can never be optimal not to know the solutions to the most frequent problems. As each worker is endowed with one efficiency unit of labor, after incurring learning costs  $1 - z/s$  efficiency units are left for production, solving a fraction  $F(z; \sigma)$  of problems drawn. Similarly, after the design cost,  $1 - z/s_K$  units of capital are left, and the machine solves a fraction  $F(z; \sigma)$  of problems drawn. Let the productivity level of an optimally trained worker of skill  $s$  in task  $(\sigma, 1)$  be denoted by  $\alpha^N(s, \sigma, 1)$ , and similarly let  $\alpha^K(s_K, \sigma)$  be the productivity level of an optimally designed machine. For simplicity, we omit the task-neutral productivity term  $A_K$  here, as it does

not affect optimal machine design. Then we have

$$\begin{aligned}\alpha^N(s, \sigma, 1) &\equiv \max_z F(z; \sigma) [1 - z/s], \\ \alpha^K(s_K, \sigma) &\equiv \max_z F(z; \sigma) [1 - z/s_K],\end{aligned}$$

A unique interior solution to the worker training and machine design problems always exists. Unlike in the baseline model, we do not require any restrictions on  $s_K$  and  $\underline{s}$  in relation to  $\bar{\sigma}$  to ensure that productivity is non-negative.<sup>36</sup> The optimal knowledge levels  $z^N(s, \sigma)$  and  $z^K(s_K, \sigma)$  are pinned down by the first-order conditions

$$\begin{aligned}f(z(s, \sigma); \sigma) [1 - z(s, \sigma)/s] &= F(z(s, \sigma, \tau); \sigma)/s, \\ f(z(s_K, \sigma); \sigma) [1 - z(s_K, \sigma)/s_K] &= F(z(s_K, \sigma); \sigma)/s_K.\end{aligned}\tag{A.1}$$

Optimality requires that the benefit of learning the solution to an additional problem—the probability that the problem occurs times the number of efficiency units left for production, be equal to the cost of doing so—the number of efficiency units lost times the fraction of problems these efficiency units would have solved. Optimal worker and machine productivities are given by

$$\alpha^N(s, \sigma, \tau) = \begin{cases} F(z(s, \sigma, \tau); \sigma) [1 - z(s, \sigma, \tau)/s] & \text{if } \tau = 1 \\ 1 & \text{if } \tau = 0 \end{cases}$$

and

$$\alpha^K(s_K, \sigma) = F(z(s_K, \sigma); \sigma) [1 - z(s_K, \sigma)/s_K].$$

Let  $\check{s}$  be an element in set  $\check{S} = s_K \cup [\underline{s}, \bar{s}]$ . By the above results, we have that  $\alpha^N(\check{s}, \sigma, 1) \equiv \alpha^K(\check{s}, \sigma)$ . Thus, workers and machines face the same productivity schedule in training-intensive tasks. We drop superscripts and define the function

$$\alpha(\check{s}, \sigma) = F(z(\check{s}, \sigma); \sigma) \left[ 1 - \frac{1}{\check{s}} z(\check{s}, \sigma) \right] \quad \check{s} \in \check{S} = s_K \cup [\underline{s}, \bar{s}],\tag{A.2}$$

where  $z(\check{s}, \sigma)$  is implicitly given by (A.1).

The qualitative properties of the productivity schedule  $\alpha(\check{s}, \sigma)$  are the same as in the baseline model. First notice that  $\alpha \in (0, 1)$  by (A.2). Furthermore, from applying the envelope theorem to (A.2) it follows that  $\alpha$  is increasing in  $\check{s}$  and decreasing in  $\sigma$ . Higher skilled factors are more productive since they face a lower learning/design cost, and productivity declines in complexity since a larger cost is incurred to achieve a given level of productivity.

To characterize comparative advantage, we again rely on log-supermodularity: Under Assumption 1, the productivity schedule  $\alpha(\check{s}, \sigma)$  is strictly log-supermodular. To show this, start by observing that  $\alpha(\check{s}, \sigma)$  is strictly log-supermodular if and only if

$$\frac{\partial^2}{\partial \check{s} \partial \sigma} \log \alpha(\check{s}, \sigma) > 0.$$

---

<sup>36</sup>A unique interior solution to the worker training problem exists if  $\tau > 0$  because first, the problem is strictly concave as  $f$  is strictly decreasing; second, the derivative of the objective at  $z = 0$  is strictly positive; finally, the value of the objective function becomes negative for a sufficiently large  $z$ . The same arguments also establish the result for the machine design problem.

Applying the envelope theorem to (A.2) yields

$$\frac{\partial}{\partial \check{s}} \log \alpha(\check{s}, \sigma) = \frac{z(\check{s}, \sigma)}{(\check{s})^2 - \check{s}z(\check{s}, \sigma)}.$$

The RHS is an increasing function of  $z(\check{s}, \sigma)$ , and so

$$\frac{\partial^2}{\partial \check{s} \partial \sigma} \log \alpha(\check{s}, \sigma) > 0 \quad \Leftrightarrow \quad \frac{\partial}{\partial \sigma} z(\check{s}, \sigma) > 0.$$

Thus,  $\alpha$  is log-supermodular if and only if optimal knowledge levels are increasing in  $\sigma$ . Differentiating the FOC (A.1) yields

$$\frac{\partial}{\partial \sigma} z(\check{s}, \sigma) = \frac{F_{\sigma} \frac{1}{\check{s}} - f_{\sigma} \left[ 1 - \frac{1}{\check{s}} z \right]}{f_z \left[ 1 - \frac{1}{\check{s}} z \right] - 2f \frac{1}{\check{s}}}.$$

The denominator of the RHS is negative as  $f_z < 0$ , and so, using the FOC we find that

$$\frac{\partial}{\partial \sigma} z(\check{s}, \sigma) > 0 \quad \Leftrightarrow \quad F f_{\sigma} > F_{\sigma} f \quad \forall Z, \sigma > 0.$$

But this condition is equivalent to  $F$  being strictly log-supermodular.

Given that  $\alpha(\check{s}, \sigma)$  has the same qualitative properties as the productivity schedule in the baseline model, assignment and comparative statics results are qualitatively the same, as well, since none of the results for the baseline model rely on the specific functional form.

## B Proofs of Formal Results Stated in the Text

### B.1 Sufficient Conditions for Existence of an Interior Equilibrium

We derive sufficient conditions ensuring that an interior equilibrium with  $\sigma_0^*, \sigma_1^* \in (\underline{\sigma}, \bar{\sigma})$  and hence  $s^* \in (\underline{s}, \bar{s})$  prevails. These conditions will consist of mild restrictions on the values that the economy's endowment of efficiency units of capital  $A_K K$  may take, given a particular choice of values  $(\bar{s}, \underline{\sigma}, \bar{\sigma})$ .

In any equilibrium in which  $k_0(\sigma) = 0$  for all  $\sigma \in [\underline{\sigma}, \bar{\sigma}]$ , we have by (8)

$$\begin{aligned} p_0(\underline{\sigma})\alpha(s_K, \underline{\sigma}) &\leq r/A_K \\ p_0(\underline{\sigma}) &= w(s^*), \end{aligned}$$

which yields  $\alpha(s_K, \underline{\sigma}) \leq r/[A_K w(s^*)]$ . Using (17) and (19) this inequality is shown to be equivalent to

$$\alpha(s_K, \underline{\sigma}) \leq \frac{\beta_1(\sigma_1^* - \underline{\sigma})}{\beta_0(\bar{\sigma} - \underline{\sigma})} \times \frac{V(s^*)}{A_K K}.$$

The RHS of the last inequality is strictly less than  $\beta_1/(\beta_0 A_K K)$ , hence a sufficient condition to rule out any equilibrium in which  $k_0(\sigma) = 0$  for all  $\sigma \in [\underline{\sigma}, \bar{\sigma}]$  is  $\alpha(s_K, \underline{\sigma}) > \beta_1/(\beta_0 A_K K)$  or

$$A_K K > \frac{\beta_1}{\beta_0} \frac{1}{\alpha(s_K, \underline{\sigma})}. \quad (\text{B.1})$$

And in any equilibrium in which  $n_1(s, \sigma) = 0$  for all  $s \in [\underline{s}, \bar{s}]$  and  $\sigma \in [\underline{\sigma}, \bar{\sigma}]$  we have by (8)

$$\begin{aligned} p_1(\bar{\sigma})\alpha(c_K, \bar{\sigma}) &= r/A_K \\ p_1(\bar{\sigma})\alpha(\bar{s}, \bar{\sigma}) &\leq w(\bar{s}) = w(s^*), \end{aligned}$$

from which we obtain  $\alpha(s_K, \bar{\sigma})/\alpha(\bar{s}, \bar{\sigma}) \geq r/[A_K w(s^*)]$ . Using (17) and (19) this inequality becomes

$$\frac{\alpha(s_K, \bar{\sigma})}{\alpha(\bar{s}, \bar{\sigma})} \geq \frac{\beta_0(\sigma_0^* - \underline{\sigma}) + \beta_1(\bar{\sigma} - \underline{\sigma})}{\beta_0(\bar{\sigma} - \sigma_0^*)} \times \frac{1}{A_K K}.$$

The RHS of the last inequality is strictly greater than  $\beta_1/(\beta_0 A_K K)$ , hence a sufficient condition to rule out any equilibrium in which  $n_1(s, \sigma) = 0$  for all  $s \in [\underline{s}, \bar{s}]$  and  $\sigma \in [\underline{\sigma}, \bar{\sigma}]$  is  $\alpha(s_K, \bar{\sigma})/\alpha(\bar{s}, \bar{\sigma}) < \beta_1/(\beta_0 A_K K)$  or

$$A_K K < \frac{\beta_1}{\beta_0} \frac{\alpha(\bar{s}, \bar{\sigma})}{\alpha(s_K, \bar{\sigma})}. \quad (\text{B.2})$$

Combining (B.1) and (B.2), we conclude that if

$$A_K K \in S, \quad S \equiv \frac{\beta_1}{\beta_0} \left( \frac{1}{\alpha(s_K, \underline{\sigma})}, \frac{\alpha(\bar{s}, \bar{\sigma})}{\alpha(s_K, \bar{\sigma})} \right),$$

then the equilibrium is interior with  $\sigma_0^*, \sigma_1^* \in (\underline{\sigma}, \bar{\sigma})$  and hence  $s^* \in (\underline{s}, \bar{s})$ . Existence of an interior equilibrium is ensured by choosing parameter values for  $(\bar{s}, \underline{\sigma}, \bar{\sigma})$  such that  $S$  is a non-empty set. Our claim that the restrictions on  $A_K K$  are mild given a particular choice of  $(\bar{s}, \underline{\sigma}, \bar{\sigma})$  is justified both for the baseline model and the more general production process in Appendix A, because we can choose parameters such that  $S \rightarrow \frac{\beta_1}{\beta_0}(1, \infty)$ . In the case of the baseline model, we can set  $\underline{\sigma} = 0$  and let  $s_K \rightarrow \bar{\sigma}$ . In the more general model, we can assume that  $\underline{\sigma}$  is sufficiently small so that  $F(Z; \underline{\sigma})$  is close to one even for very small  $Z$ ; and that  $\bar{\sigma}$  is sufficiently large so that  $F(Z; \bar{\sigma})$  is close to zero even for very large  $Z$ , while at the same time  $\bar{s}$  is large so that  $\alpha(\bar{s}, \bar{\sigma})$  stays finite.



## B.2 Proofs of Lemmas Stated in the Text

**Proof of Lemma 1** Follows from the definition of strict log-supermodularity and simple differentiation. ■

**Proof of Lemma 2** (a) For any vectors  $(s, \sigma)$  and  $(s', \sigma')$  such that  $n_0(s, \sigma) > 0$  and  $n_1(s', \sigma') > 0$  we have by the zero-profit condition (8)  $p_0(\sigma) = w(s)$  and  $p_0(\sigma) \leq w(s')$ , or  $w(s) \leq w(s')$ , and

$$\begin{aligned} p_1(\sigma')\alpha(s', \sigma') &= w(s'), \\ p_1(\sigma')\alpha(s, \sigma') &\leq w(s). \end{aligned}$$

Together these conditions imply  $\alpha(s', \sigma')/\alpha(s, \sigma') \geq 1$ . Since  $\alpha$  is increasing in  $s$  we must have  $s' \geq s$ . Furthermore, it must be that  $s^* > \underline{s}$ , for suppose not. Then market clearing (3) implies that  $k_0(\sigma) > 0$  for all  $\sigma$  (task output must be strictly positive due to the INADA properties of the Cobb-Douglas production function). By (8), for some  $(s, \sigma)$

$$\begin{aligned} p_1(\sigma)\alpha(s, \sigma) &= w(s), \\ p_1(\sigma)\alpha(s_K, \sigma) &\leq r/A_K, \end{aligned}$$

which yields

$$\frac{w(s)}{r/A_K} \leq \frac{\alpha(s, \sigma)}{\alpha(s_K, \sigma)}.$$

Furthermore,  $p_0(\sigma)\alpha(s_K, \sigma) = r/A_K$  and  $p_0(\sigma) \leq w(s)$ . This yields

$$\frac{w(s)}{r/A_K} \geq \frac{1}{\alpha(s_K, \sigma)}.$$

Together with the previous result this implies  $\alpha(s, \sigma) \geq 1$  which is impossible given (9).

(b) If  $k_0(\sigma) > 0$ , then by the zero-profit condition (8)

$$\frac{w(s^*)}{r/A_K} \geq \frac{1}{\alpha(s_K, \sigma)},$$

and there is some  $\sigma'$  such that  $n_1(s^*, \sigma') > 0$  and hence by (8)

$$\frac{w(s^*)}{r/A_K} \leq \frac{\alpha(s^*, \sigma')}{\alpha(s_K, \sigma')}.$$

The previous two inequalities imply

$$\frac{\alpha(s^*, \sigma')}{\alpha(s_K, \sigma')} \geq \frac{1}{\alpha(s_K, \sigma)},$$

but since  $\alpha(s_K, \sigma) < 1$ , we have  $\alpha(s^*, \sigma')/\alpha(s_K, \sigma') > 1$  which is only possible if  $s^* > s_K$ .

Next, observe that for any  $(\sigma, \sigma')$  and  $s \leq s^*$  such that  $k_0(\sigma) > 0$  and  $n_0(s, \sigma') > 0$  we have by (8),

$$\begin{aligned} p_0(\sigma)\alpha(s_K, \sigma) &= r/A_K \\ p_0(\sigma) &\leq w(s), \end{aligned}$$

and

$$\begin{aligned} p_0(\sigma')\alpha(s_K, \sigma') &\leq r/A_K \\ p_0(\sigma') &= w(s), \end{aligned}$$

which yields  $\alpha(s_K, \sigma) \geq \alpha(s_K, \sigma')$  and so  $\sigma \leq \sigma'$ . Thus we have established existence of  $\sigma_0^*$ .

Similarly, for any  $(\sigma, \sigma')$  and  $s \geq s^*$  such that  $k_1(\sigma) > 0$  and  $n_1(s, \sigma') > 0$ , we have by (8),

$$\begin{aligned} p_1(\sigma)\alpha(s_K, \sigma) &= r/A_K \\ p_1(\sigma)\alpha(s, \sigma) &\leq w(s), \end{aligned}$$

and

$$\begin{aligned} p_1(\sigma')\alpha(s_K, \sigma') &\leq r/A_K \\ p_1(\sigma')\alpha(s, \sigma') &= w(s), \end{aligned}$$

which yields

$$\frac{\alpha(s_K, \sigma)}{\alpha(s, \sigma)} \geq \frac{\alpha(s_K, \sigma')}{\alpha(s, \sigma')},$$

and so  $\sigma \leq \sigma'$  by the log-supermodularity of  $\alpha$  and since  $s > s_K$ . This establishes existence of  $\sigma_1^*$ .

Now, it must be that  $\sigma_0^* < \sigma_1^*$ , for suppose not. If  $\sigma_0^* > \sigma_1^*$ , then there exist  $(s, \sigma)$  such that  $k_0(\sigma) > 0$ ,  $k_1(\sigma) = 0$ ,  $n_0(s, \sigma) = 0$ , and  $n_1(s, \sigma) > 0$ . By (8),

$$\begin{aligned} p_0(\sigma)\alpha(s_K, \sigma) &= r/A_K \\ p_0(\sigma) &\leq w(s), \end{aligned}$$

and

$$\begin{aligned} p_1(\sigma)\alpha(s_K, \sigma) &\leq r/A_K \\ p_1(\sigma)\alpha(s, \sigma) &= w(s). \end{aligned}$$

This yields  $\alpha(s, \sigma) \geq 1$  which contradicts (9). If  $\sigma_0^* = \sigma_1^*$ , then similar arguments lead to  $\alpha(s, \sigma) = 1$ , which also contradicts (9). ■

**Proof of Lemma 3** Given Lemma 2, the problem is to match workers of skill levels  $s \in [s^*, \bar{s}]$  to tasks  $\sigma \in [\sigma_1^*, \bar{\sigma}]$  in a setting identical to that in Costinot and Vogel (2010). Hence, the proof of Lemma 1 from their paper applies. ■

### B.3 Proofs of Propositions Stated in the Text

**Proof of Proposition 1** We first show that in the absence of changes to the distribution of skills, a flattening (steepening) of the matching function at the upper end implies an upward (downward) shift of the matching function everywhere. Formally, if  $\widehat{M}'(\bar{s}) < M'(\bar{s})$ , then  $\widehat{M}(s) < M(s)$  for all  $s \in [\max\{s^*, \widehat{s}^*\}, \bar{s}]$ . For suppose that  $\widehat{M}'(\bar{s}) < M'(\bar{s})$  and that there exists some  $s' \in [\max\{s^*, \widehat{s}^*\}, \bar{s})$  such that  $\widehat{M}(s') \leq M(s')$ . Then there exists some  $s'' \in [s', \bar{s})$  such that  $\widehat{M}(s'') = M(s'')$ ,  $\widehat{M}'(s'') \geq M'(s'')$ , and  $\widehat{M}(s) > M(s)$  for all  $s \in (s'', \bar{s})$ . We will show that this leads to a contradiction.

Integrating (11) yields an expression for the wage premium of the most skilled worker with respect to any other skill group employed in training-intensive tasks,

$$\frac{w(\bar{s})}{w(s)} = \omega(s; M), \quad s \geq s^*$$

where

$$\omega(s; M) \equiv \exp \left[ \int_s^{\bar{s}} \frac{\partial}{\partial z} \log \alpha(z, M(z)) dz \right]. \quad (\text{B.3})$$

Because  $\alpha$  is increasing in its first argument,  $\omega$  is decreasing in  $s$ . Moreover, by the log-supermodularity of  $\alpha$ , if  $\widehat{M}(z) > M(z)$  for all  $z \in (s, \bar{s})$  and any  $s$  that belongs to the domains of both  $\widehat{M}$  and  $M$ , then  $\omega(s; \widehat{M}) > \omega(s; M)$ .

Plugging (B.3) into (10), we obtain

$$\frac{M'(\bar{s})}{M'(s)} = \omega(s; M) \frac{v(\bar{s})}{v(s)}. \quad (\text{B.4})$$

Therefore,

$$\frac{\widehat{M}'(\bar{s})}{M'(\bar{s})} = \frac{\omega(s''; \widehat{M}) \widehat{M}'(s'')}{\omega(s''; M) M'(s'')}.$$

By the above arguments, the right side of the last equation is larger than one, so that we must have  $\widehat{M}'(\bar{s}) > M'(\bar{s})$ , a contradiction. A similar argument establishes that a steepening at the upper end leads to a downward shift everywhere.

*Proof that  $\widehat{\sigma}_1^* > \sigma_1^*$*  First suppose  $\widehat{\sigma}_1^* \leq \sigma_1^*$  and  $\widehat{M}'(\bar{s}) \geq M'(\bar{s})$ .

By (22) and (B.4),

$$\frac{V(s^*)}{\bar{\sigma} - \sigma_0^*} \times \frac{M'(\bar{s})}{\omega(s^*; M)} = \frac{\beta_0 v(\bar{s})}{\beta_1}. \quad (\text{B.5})$$

This together with (20), implies

$$\frac{A_K \alpha(s_K, \sigma_0^*) K}{\beta_0(\sigma_0^* - \underline{\sigma}) + \beta_1(\sigma_1^* - \underline{\sigma})} \times \frac{M'(\bar{s})}{\omega(s^*; M)} = \frac{v(\bar{s})}{\beta_1}. \quad (\text{B.6})$$

Suppose that  $\widehat{s}^* \geq s^*$ . Then (B.5) implies that  $\widehat{\sigma}_0^* < \sigma_0^*$ , while (B.6) implies  $\widehat{\sigma}_0^* > \sigma_0^*$ , a contradiction. So we must have  $\widehat{s}^* < s^*$ . If  $\widehat{\sigma}_0^* \geq \sigma_0^*$ , then from (21),  $\widehat{s}^* > s^*$ ,<sup>37</sup> so it must be that  $\widehat{\sigma}_0^* < \sigma_0^*$ . Then by 21,  $\alpha(\widehat{s}_K, \widehat{\sigma}_0^*) > \alpha(s_K, \sigma_0^*)$ . This implies that the LHS of (20) increases, while the RHS decreases, a contradiction.

Next, suppose that  $\widehat{\sigma}_1^* \leq \sigma_1^*$  and  $\widehat{M}'(\bar{s}) < M'(\bar{s})$ . We have shown that in this case the matching function shifts up, so we must have  $\widehat{s}^* \leq s^*$ . Then  $\widehat{\sigma}_0^* < \sigma_0^*$  from (21). But we have just shown that it is impossible to have  $\widehat{\sigma}_1^* \leq \sigma_1^*$ ,  $\widehat{\sigma}_0^* < \sigma_0^*$ , and  $\widehat{s}^* \leq s^*$  at the same time. Thus we have established that  $\widehat{\sigma}_1^* > \sigma_1^*$ .

*Proof that  $\widehat{M}(s) > M(s)$*  Suppose that  $\widehat{M}'(\bar{s}) > M'(\bar{s})$ , which we have shown implies  $\widehat{M}(s) < M(s)$  and, by (B.4),  $\widehat{M}'(s) > M'(s)$  for all  $s$  belonging to the domains of both  $\widehat{M}$  and  $M$ . As we have established that  $\widehat{\sigma}_1^* < \sigma_1^*$ , by the properties of the matching function we must have  $\widehat{s}^* > s^*$ . By (10), the wage share of a worker who is always assigned to training-intensive tasks has increased,

$$\frac{\widehat{w}(s)}{\widehat{Y}} = \frac{\beta_1 \widehat{M}'(s)}{\mu v(s)} > \frac{\beta_1 M'(s)}{\mu v(s)} = \frac{w(s)}{Y} \quad \forall s \in [\widehat{s}^*, \bar{s}].$$

But this means that the wage shares of all remaining workers have increased, as well,

$$\frac{\widehat{w}(s)}{\widehat{Y}} = \frac{\widehat{w}(\widehat{s}^*)}{\widehat{Y}} > \frac{w(\widehat{s}^*)}{Y} > \frac{w(s)}{Y} \quad \forall s \in [\underline{s}, \widehat{s}^*),$$

---

<sup>37</sup>To see this, rewrite (21) as

$$\frac{\alpha(s_K, \sigma_1^*)}{\alpha(s_K, \sigma_0^*) \alpha(s^*, \sigma_1^*)} = 1.$$

By the log-supermodularity of  $\alpha$ , a rise in  $s_K$  leads the ratio  $\alpha(s_K, \sigma_1^*)/\alpha(s_K, \sigma_0^*)$  to rise since  $\sigma_1^* > \sigma_0^*$ . Again due to log-supermodularity, the fall in  $\sigma_1^*$  raises the ratio  $\alpha(s_K, \sigma_1^*)/\alpha(s^*, \sigma_1^*)$  since  $s_K < s^*$ . The rise in  $\sigma_0^*$  raises the LHS further. Therefore,  $s^*$  must increase.

where the last inequality is due to (23). Therefore, the total labor share has increased,

$$\frac{\int_{\bar{s}}^{\bar{s}} \widehat{w}(s)v(s)ds}{\widehat{Y}} > \frac{\int_{\bar{s}}^{\bar{s}} w(s)v(s)ds}{Y}.$$

By (10) and (19), this implies  $\beta_0\widehat{\sigma}_0^* + \beta_1\widehat{\sigma}_1^* < \beta_0\sigma_0^* + \beta_1\sigma_1^*$ .

Now observe that if  $\widehat{M}(s) < M(s)$  then  $\omega(\widehat{s}^*; \widehat{M}) < \omega(s^*; M)$  since also  $\widehat{s}^* > s^*$ . By (B.5), we must have  $\widehat{\sigma}_0^* < \sigma_0^*$ . But this means that (B.6) can only hold if also the total labor share has decreased,  $\beta_0\widehat{\sigma}_0^* + \beta_1\widehat{\sigma}_1^* > \beta_0\sigma_0^* + \beta_1\sigma_1^*$ , a contradiction.

*Proof that if  $\widehat{s}_K \geq s^*$  then  $\widehat{s}^* > s^*$*  Immediate from Lemma 2 which says that  $\widehat{s}^* > \widehat{s}_K$ . ■

**Proof of Proposition 2** We proceed in three steps.

1. If the labor share increases, then the marginal training-intensive task becomes less complex. Formally, if  $\beta_0\widehat{\sigma}_0^* + \beta_1\widehat{\sigma}_1^* < \beta_0\sigma_0^* + \beta_1\sigma_1^*$ , then  $\widehat{\sigma}_1^* < \sigma_1^*$ . For suppose that  $\beta_0\widehat{\sigma}_0^* + \beta_1\widehat{\sigma}_1^* < \beta_0\sigma_0^* + \beta_1\sigma_1^*$ , but  $\widehat{\sigma}_1^* \geq \sigma_1^*$ . Then  $\widehat{\sigma}_0^* < \sigma_0^*$ . By (21),  $\widehat{s}^* < s^*$ . But by (20),  $\widehat{s}^* > s^*$ , a contradiction.
2. If the marginal training-intensive task becomes less complex, then the marginal worker becomes more skilled. Formally, if  $\widehat{\sigma}_1^* < \sigma_1^*$ , then  $\widehat{s}^* > s^*$ . For suppose that  $\widehat{\sigma}_1^* < \sigma_1^*$  but  $\widehat{s}^* \leq s^*$ . Then (21) implies  $\widehat{\sigma}_0^* < \sigma_0^*$ . But since  $\widehat{V}(\widehat{s}^*) < V(s^*)$ , (20) implies  $\widehat{\sigma}_0^* > \sigma_0^*$ , a contradiction.
3. If at one point the new matching function is flatter and does not lie below the old matching function, then it lies above the old one everywhere to the left of this point. Formally, if  $\widehat{M}'(s') \leq M'(s')$  and  $\widehat{M}(s') \geq M(s')$  for some  $s' \in (\max\{s^*, \widehat{s}^*\}, \bar{s}]$ , then  $\widehat{M}(s) \geq M(s)$  for all  $s \in [\max\{s^*, \widehat{s}^*\}, s']$ . For suppose that  $\widehat{M}'(s') \leq M'(s')$  and  $\widehat{M}(s') \geq M(s')$ , and that there exists some  $s'' \in [\max\{s^*, \widehat{s}^*\}, s')$  such that  $\widehat{M}(s'') < M(s'')$ . Then there exists some  $s''' \in (s'', s')$  such that  $\widehat{M}(s''') = M(s''')$ ,  $\widehat{M}'(s''') > M'(s''')$ , and  $\widehat{M}(s) \geq M(s)$  for all  $s \in [s''', s']$ . By (10),

$$\frac{\widehat{M}'(s''')}{M'(s''')} = \frac{\widehat{w}(s''')/\widehat{w}(s')}{w(s''')/w(s')} \times \frac{\widehat{v}(s''')/\widehat{v}(s')}{v(s''')/v(s')} \times \frac{\widehat{M}'(s')}{M'(s')}.$$

Since  $\widehat{V} \succeq V$ , and because the upward shift of the matching function raises inequality and thus lowers the wage of type  $s'''$  relative to that of type  $s'$ , the right side of the last equation is no greater than one, so that  $\widehat{M}'(s''') \leq M'(s''')$ , a contradiction.

Thus, we have shown that if the increase in skill abundance results in an increase in the labor share, then the lower endpoint of the matching function moves southeast (Steps 1 and 2). This means that the matching function must shift down everywhere, for if it shifted up at one point, it would shift up everywhere (Step 3), and it would be impossible for its lower endpoint to move southeast. ■

## B.4 Proofs of Corollaries Stated in the Text

**Proof of Corollary 1** Integrating (11), the first part of the result is immediate given the shift in the matching function and the log-supermodularity of  $\alpha$ . The second part follows since  $\widehat{w}(s')/\widehat{w}(s) = 1$  but  $w(s')/w(s) > 1$  for all such  $s', s$ . ■

**Proof of Corollary 2** Recall that the labor share is proportional to  $\beta_0(\bar{\sigma} - \sigma_0^*) + \beta_1(\bar{\sigma} - \sigma_1^*)$ . As  $\hat{\sigma}_1^* > \sigma_1^*$ , the result is immediate if  $\hat{\sigma}_0^* \geq \sigma_0^*$ . Then consider the case  $\hat{\sigma}_0^* < \sigma_0^*$ . Rewrite (20) as

$$A_K \alpha(s_K, \sigma_0^*) K = \frac{\beta_0(\sigma_0^* - \underline{\sigma}) + \beta_1(\sigma_1^* - \underline{\sigma})}{\frac{\beta_0(\bar{\sigma} - \sigma_0^*)}{V(s^*)}}.$$

The LHS increases. If the denominator of the RHS increases, then so must the numerator, which is proportional to the capital share. Hence the labor share decreases. If the denominator of the RHS decreases, then the wage share of all workers falls, again implying a fall in the labor share.

■

**Proof of Corollary 3** Analogous to the proof of Corollary 1. ■

## C A Model with Fixed Costs

### C.1 Model setup

Worker training technologies are as in the baseline model. However, we now assume that an upfront expense of  $\varphi(\sigma)$  is required to equip the firm's stock of machines with  $\sigma$  units of knowledge. This cost is independent of the size of the stock, as in the case of software. We make the critical assumption  $\varphi' > 0$ , and for simplicity we set  $\varphi(\underline{\sigma}) = 0$  and  $\varphi'' > 0$ . In our numerical solutions we choose  $\varphi(\sigma) = c_K(\sigma - \underline{\sigma})^2$  where  $c_K$  is the parameter capturing labor-replacing technical change. Machines capable of performing a task produce  $A_K$  units of task output—that is, machine productivity is independent of complexity. As in the baseline model, worker productivity is independent of complexity and worker skill in innate ability tasks, but in training-intensive tasks worker productivity is given by  $\alpha(s, \sigma) \equiv 1 - \sigma/s$ .

We assume that each task is produced by a single monopolistic firm.<sup>38</sup> In contrast, final good firms are perfectly competitive just as in the baseline version of the model. The final good production function is now

$$Y = \left[ \int_{\underline{\sigma}}^{\bar{\sigma}} \left\{ \beta_1 y_0(\sigma)^{\frac{\varepsilon-1}{\varepsilon}} + \beta_1 y_1(\sigma)^{\frac{\varepsilon-1}{\varepsilon}} \right\} d\sigma \right]^{\frac{\varepsilon}{\varepsilon-1}}, \quad (\text{C.1})$$

with  $\varepsilon > 1$  and  $\sum_{\tau} \beta_{\tau} = 1$  for CRS. Given profit maximization by final good firms, the CES production function yields the standard isoelastic input demand curve, inducing the well-known constant-markup pricing rule.

Standard arguments establish that equilibrium variable profits of the firm supplying task  $(\sigma, \tau)$  are given by

$$\pi_{\tau}(\sigma, \chi) = \beta_{\tau}^{\varepsilon} \frac{(\varepsilon - 1)^{\varepsilon-1}}{\varepsilon^{\varepsilon}} \chi^{-(\varepsilon-1)} Y \quad (\text{C.2})$$

where  $\chi$  is marginal cost which depends on the characteristics of tasks and factors employed. In particular, if employing labor of type  $s$  we have

$$\chi \equiv \chi(s, \sigma, \tau) = \begin{cases} w(s) & \text{if } \tau = 0 \\ w(s)/\alpha(s, \sigma) & \text{if } \tau = 1, \end{cases}$$

and if employing capital,

$$\chi = \frac{r}{A_K}.$$

Furthermore, equilibrium task output is

$$y_{\tau}(\sigma, \chi) = \beta_{\tau}^{\varepsilon} \left( \frac{\varepsilon - 1}{\varepsilon} \right)^{\varepsilon} \chi^{-\varepsilon} Y. \quad (\text{C.3})$$

### C.2 Equilibrium assignment

We focus on equilibria in which  $r/A_K < w(\bar{s})$ , so that the marginal cost of using machines is less than that of employing labor of any type in any task.<sup>39</sup> In such an equilibrium, the assignment is qualitatively the same as in the case we analyzed for the baseline model. Task producers employ the factor that delivers the highest total profits. Note that variable profits among firms that use

<sup>38</sup>Holmes and Mitchell (2008) present a more complex model where labor and machines are optimally assigned to tasks *within* monopolistic firms. We suspect that our results would hold in a version of that model as well.

<sup>39</sup>We do not derive formal conditions ensuring this property, but it will feature in our numerical solutions.

machines are constant across tasks. In any equilibrium with the above characteristic there are a threshold tasks  $\sigma_0^*$ ,  $\sigma_1^*$  such that it is optimal for firms to use machines in all innate ability tasks with  $\sigma \leq \sigma_0^*$  and in all training-intensive tasks with  $\sigma \leq \sigma_1^*$ .<sup>40</sup> As in the baseline model, there is a cutoff  $s^*$  such that workers below the cutoff perform innate ability tasks, while workers above the cutoff carry out training intensive tasks, with higher skilled workers performing more complex tasks.

### C.3 Solving the model

The threshold tasks are determined by no-arbitrage conditions. Profits from employing labor in these tasks must be equal to profits from using machines. This means that the difference in variable profits between machines and labor must equal the fixed cost of designing the machine,

$$\beta_0^\varepsilon \frac{(\varepsilon - 1)^{\varepsilon-1}}{\varepsilon^\varepsilon} Y \left[ \left( \frac{A_K}{r} \right)^{\varepsilon-1} - \left( \frac{1}{w(s^*)} \right)^{\varepsilon-1} \right] = \varphi(\sigma_0^*) \quad (\text{C.4})$$

and

$$\beta_1^\varepsilon \frac{(\varepsilon - 1)^{\varepsilon-1}}{\varepsilon^\varepsilon} Y \left[ \left( \frac{A_K}{r} \right)^{\varepsilon-1} - \left( \frac{\alpha(s^*, \sigma_1^*)}{w(s^*)} \right)^{\varepsilon-1} \right] = \varphi(\sigma_1^*). \quad (\text{C.5})$$

Worker assignment to training intensive tasks is assortative as in the baseline model, with the wage schedule given by

$$\frac{d \log w(s)}{ds} = \frac{\partial \log \alpha(s, M(s))}{\partial s} \quad (\text{C.6})$$

and the matching function satisfying

$$M'(s) = \frac{\alpha(s, M(s))^{1-\varepsilon} v(s) w(s)^\varepsilon}{\beta_1^\varepsilon \left( \frac{\varepsilon-1}{\varepsilon} \right)^\varepsilon Y} \quad (\text{C.7})$$

as well as the boundary conditions  $M(s^*) = \sigma_1^*$  and  $M(\bar{s}) = \bar{\sigma}$ .

The model is closed by market clearing conditions and the production function. For capital markets to clear, we must have  $\int_{\underline{\sigma}}^{\sigma_0^*} k_0(\sigma) d\sigma + \int_{\underline{\sigma}}^{\sigma_1^*} k_1(\sigma) d\sigma = K$ . Because of our assumption of constant marginal costs of using machines regardless of a task's complexity, we have that task outputs are constant within innate ability tasks and within training-intensive tasks, and so are machine inputs. Thus, the capital market clearing condition becomes  $(\sigma_0^* - \underline{\sigma})k_0 + (\sigma_1^* - \underline{\sigma})k_1 = K$ . Using the task production function  $y_\tau = A_K k_\tau$  and equilibrium task output (C.3), we obtain  $k_0/k_1 = (\beta_0/\beta_1)^\varepsilon$ . Together with the market clearing condition and the task production function this implies

$$y_\tau(\sigma) = \frac{\beta_\tau^\varepsilon A_K K}{\sum_\tau \beta_\tau (\sigma_\tau^* - \underline{\sigma})} \quad \text{for all } \sigma \in [\underline{\sigma}, \sigma_\tau^*]. \quad (\text{C.8})$$

When firms employ labor in innate ability tasks, the task production function is  $y_0(\sigma) =$

---

<sup>40</sup>Notice that profits of firms employing type- $s$  labor are flat in  $\sigma$  ( $\tau = 0$ ) or a decreasing and convex function of  $\sigma$  approaching zero as  $\sigma$  gets large ( $\tau = 1$ ). However, because fixed machine design costs are decreasing in  $\sigma$ , and because of the convexity of the cost function, profits of firms using machines become negative as  $\sigma$  gets large. Due to the properties of  $\alpha$ , profits when employing labor are a convex function of  $\sigma$ , so that there exists a value of  $\sigma$  such that profits from employing type- $s$  labor are equal to profits from using machines.

Table C.1: Parameter values for the model with fixed design costs

$\beta_0$	=	1/3
$\varepsilon$	=	2
$\sigma$	∈	[0, 1]
$s$	∈	[1.01, 2]
$v(s)$		see text
$A_K K$	=	1
$\phi(\sigma)$	=	$c_K \sigma^2$
$c_K$	∈	[1, 2]

$\int_{\underline{s}}^{s^*} n_0(s, \sigma) ds$ . The market clearing condition is  $\int_{\sigma_0^*}^{\bar{\sigma}} \int_{\underline{s}}^{s^*} n_0(s, \sigma) ds d\sigma = V(s^*)$ , and so

$$y_0(\sigma) = \frac{V(s^*)}{\bar{\sigma} - \sigma_0^*} \quad \text{for all } \sigma \in [\sigma_0^*, \bar{\sigma}]. \quad (\text{C.9})$$

Given (12), (C.8), and (C.9), final good output must satisfy

$$\begin{aligned} Y^{\frac{\varepsilon-1}{\varepsilon}} &= [\beta_0^\varepsilon(\sigma_0^* - \underline{\sigma}) + \beta_1^\varepsilon(\sigma_1^* - \underline{\sigma})]^{\frac{1}{\varepsilon}} (A_K K)^{\frac{\varepsilon-1}{\varepsilon}} \\ &\quad + \beta_0 [\bar{\sigma} - \sigma_0^*]^{\frac{1}{\varepsilon}} V(s^*)^{\frac{\varepsilon-1}{\varepsilon}} + \beta_1 \int_{s^*}^{\bar{s}} M'(s)^{\frac{1}{\varepsilon}} [\alpha(s, M(s))v(s)]^{\frac{\varepsilon-1}{\varepsilon}} ds. \end{aligned} \quad (\text{C.10})$$

Plugging (C.8) into (C.3) yields an expression for the rental rate,

$$\frac{r}{A_K} = \frac{\varepsilon - 1}{\varepsilon} [\beta_0^\varepsilon(\sigma_0^* - \underline{\sigma}) + \beta_1^\varepsilon(\sigma_1^* - \underline{\sigma})]^{\frac{1}{\varepsilon}} \left( \frac{Y}{A_K K} \right)^{\frac{1}{\varepsilon}}. \quad (\text{C.11})$$

Similarly, plugging (C.9) into (C.3) gives an expression for the wage paid to the marginal worker,

$$w(s^*) = \frac{\varepsilon - 1}{\varepsilon} \beta_0 (\bar{\sigma} - \sigma_0^*)^{\frac{1}{\varepsilon}} \left( \frac{Y}{V(s^*)} \right)^{\frac{1}{\varepsilon}}. \quad (\text{C.12})$$

We solve the model by grid search. Given a guess of  $(s^*, \sigma_1^*)$ , we solve for the matching function and the wage distribution, obtaining  $w(s^*)/Y^{1/\varepsilon}$ . Using this, we calculate  $\sigma_0^*$  from (C.12) and  $Y$  from (C.10), thus obtaining  $w(s^*)$ . We then calculate  $r/A_K$  from (C.11). Finally, we check whether the no-arbitrage conditions (C.4) and (C.5) are satisfied.

#### C.4 Numerical solution

We solve the model for values of  $c_K$  ranging from one to two. Parameter values are given in Table C.1. The skill distribution  $v(s)$  is a truncated log-normal. We construct the distribution such that the corresponding (non-truncated) normal distribution has mean  $\log \underline{s} + 1/3 * \log(\bar{s}/\underline{s})$  and standard deviation  $1/6 * \log(\bar{s}/\underline{s})$ . This implies that the original normal distribution is truncated at four (two) standard deviations above (below) the mean.

As in the baseline model, firms adopt machines more widely in training-intensive tasks as design becomes cheaper—the marginal training-intensive task becomes more complex. The effect on the marginal innate ability task is ambiguous, however. (See Figure C.1.)

The skill cutoff increases as design gets cheaper, so that the employment share of innate



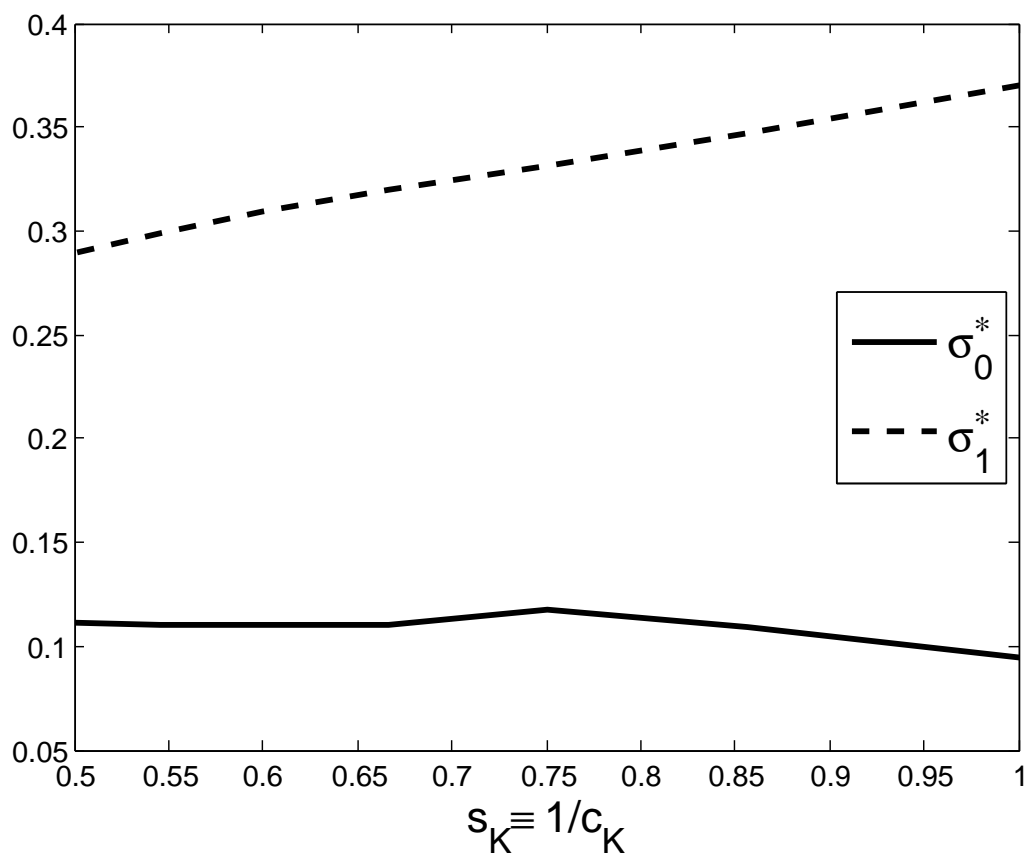


Figure C.1: Changes in cutoff tasks in the model with fixed design cost as machine design becomes cheaper.

ability tasks rises. (See Figure C.2.) But because  $\sigma_1^*$  also increases the matching function shifts up, implying a reallocation of workers to more complex training-intensive tasks. Market clearing implies a compression of the wage distribution in the lower part of the wage distribution but increasing dispersion in the upper part. Thus, the model with a fixed design cost features job and wage polarization just like the baseline model.

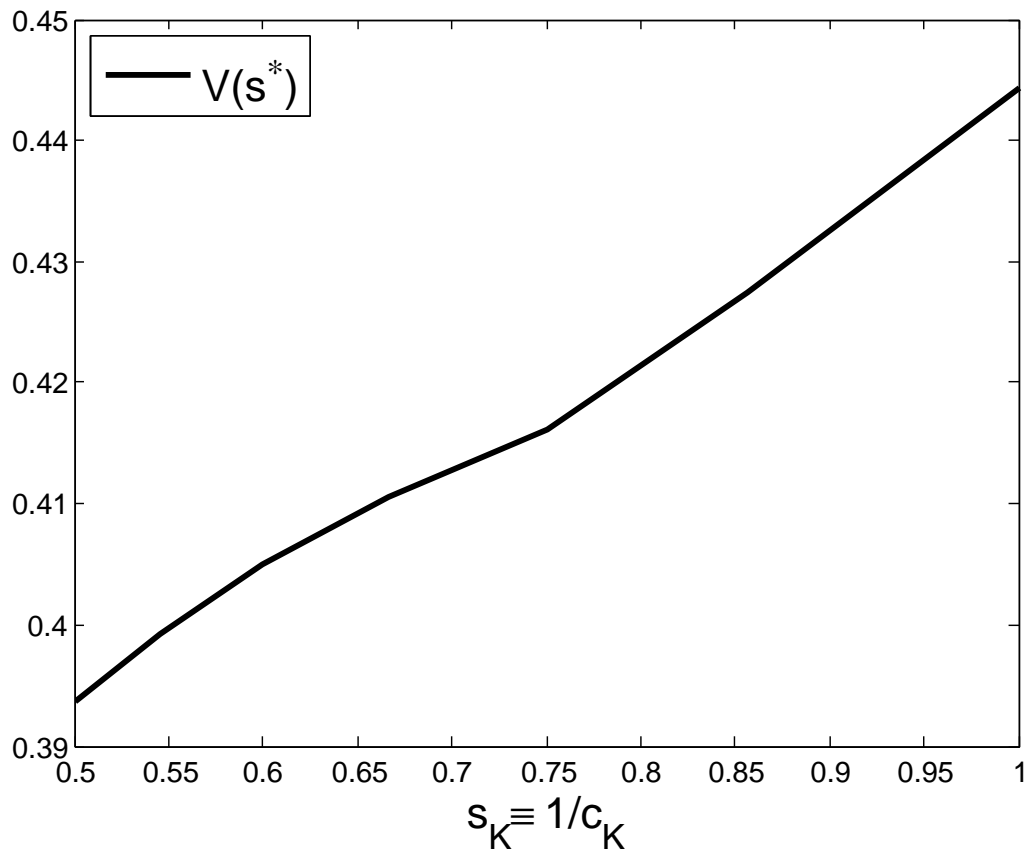


Figure C.2: Changes in the skill cutoff in the model with fixed design cost as machine design becomes cheaper. The fraction of workers below the cutoff  $V(s^*)$  is plotted.

Table D.2: Measuring Training Requirements Based on SVP and Job Zones

	<i>SVP</i>	<i>Job Zone</i>	<i>Training time</i>
1	short demonstration	1	1.5 months
2	up to 30 days	1	1.5 months
3	30 days to 3 months	1	1.5 months
4	3 to 6 months	2	7.5 months
5	6 months to 1 year	2	7.5 months
6	1 to 2 years	3	1.5 years
7	2 to 4 years	4	3 years
8	4 to 10 years	5	7.5 years
9	over 10 years	5	7.5 years

## D Data Sources and Measurement of Training Requirements

*Data sources.*—Our 1971 training measure comes from the Fourth Edition Dictionary of Occupational Titles (DOT), which is made available in combination with the 1971 April Current Population Survey (CPS) (National Academy of Sciences 1981). We obtain contemporaneous wage data from the IPUMS 1970 census extract (the processing of this data follows the procedure of Acemoglu and Autor (2011)). Our 2007 training measure comes from the Job Zones file in the O\*NET database available at <http://www.onetcenter.org/database.html?p=2>. For contemporaneous micro data we use the IPUMS 2008 American Community Survey (ACS).

*Measuring training requirements.*—SVP (see definition in Section 5.2) is measured on a nine-point scale in the DOT. In the O\*NET database, Job Zones are measured on a five-point scale which maps into the nine-point SVP scale. See Table D.2 for the interpretation of the SVP scale and the mapping between SVP and Job Zones. In the DOT data, we convert SVP into Job Zones. We assign midpoints to consistently measure training requirements over time. We assign a conservative value to the highest category. See the last column in Table D.2 for details.

The DOT variables, including SVP, in the 1971 April CPS extract vary at the level of 4,528 distinct occupations. For the occupation-level analysis, we collapse the CPS micro data to the three-digit occupation level using David Dorn’s classification of occupations (Dorn 2009), weighting by the product of sampling weights and hours worked. The Job Zones variable in the O\*NET database is available for 904 distinct occupations of the Standard Occupational Classification System (SOC). In the 2008 ACS data there are 443 distinct SOC occupations. We collapse the O\*NET data to these 443 occupations and then merge it to the ACS data. For the occupation-level analysis, we collapse the ACS micro data to the three-digit occupation level in the same way as the CPS data.

Table D.3 lists the twenty least and most training-intensive occupations (using David Dorn’s classification) in 1971. Table D.4 does the same for 2007. Table D.5 lists the twenty occupations experiencing the largest declines and increases in training requirements.

Table D.3: Least and Most Training-Intensive Occupations, 1971

Occupation (occ1990dd grouping)	Training requirements in years (1971)
<i>a) least training-intensive</i>	
Public transportation attendants and inspectors	0.1
Packers and packagers by hand	0.2
Waiter/waitress	0.2
Mail carriers for postal service	0.3
Garage and service station related occupations	0.4
Bartenders	0.4
Messengers	0.4
Parking lot attendants	0.4
Cashiers	0.5
Child care workers	0.6
Misc material moving occupations	0.6
Taxi cab drivers and chauffeurs	0.7
Baggage porters	0.7
Housekeepers, maids, butlers, stewards, and lodging quarters cleaners	0.7
Typists	0.7
Mail and paper handlers	0.7
Proofreaders	0.7
Bus drivers	0.7
File clerks	0.7
Helpers, surveyors	0.8
<i>b) most training-intensive</i>	
Musician or composer	6.8
Mechanical engineers	6.8
Aerospace engineer	6.8
Electrical engineer	6.9
Biological scientists	6.9
Chemical engineers	7.0
Chemists	7.0
Managers in education and related fields	7.0
Petroleum, mining, and geological engineers	7.1
Architects	7.1
Subject instructors (HS/college)	7.1
Dentists	7.2
Veterinarians	7.2
Lawyers	7.2
Civil engineers	7.2
Clergy and religious workers	7.3
Psychologists	7.3
Physicians	7.3
Geologists	7.5
Physicists and astronomers	7.5

Table D.4: Least and Most Training-Intensive Occupations, 2007

Occupation (occ1990dd grouping)	Training requirements in years (2007)
<i>a) least training-intensive</i>	
Waiter/waitress	0.1
Misc food prep workers	0.1
Ushers	0.1
Parking lot attendants	0.1
Kitchen workers	0.1
Furniture and wood finishers	0.1
Pressing machine operators (clothing)	0.1
Fishers, hunters, and kindred	0.1
Textile sewing machine operators	0.1
Graders and sorters of agricultural products	0.1
Garage and service station related occupations	0.1
Taxi cab drivers and chauffeurs	0.1
Animal caretakers, except farm	0.2
Butchers and meat cutters	0.3
Janitors	0.4
Sales demonstrators / promoters / models	0.4
Housekeepers, maids, butlers, stewards, and lodging quarters cleaners	0.4
Miners	0.4
Cashiers	0.4
Stock and inventory clerks	0.4
<i>b) most training-intensive</i>	
Other health and therapy	7.5
Psychologists	7.5
Physicians	7.5
Economists, market researchers, and survey researchers	7.5
Lawyers	7.5
Managers of medicine and health occupations	7.5
Physicians' assistants	7.5
Biological scientists	7.5
Medical scientists	7.5
Physical scientists, n.e.c.	7.5
Podiatrists	7.5
Veterinarians	7.5
Subject instructors (HS/college)	7.5
Dietitians and nutritionists	7.5
Urban and regional planners	7.5
Pharmacists	7.5
Librarians	7.5
Optometrists	7.5
Dentists	7.5
Physicists and astronomers	7.5

Table D.5: Largest Decreases and Increases in Training Requirements, 1971-2007

Occupation (occ1990dd grouping)	Change in training requirements (years) 1971-2007	Training requirements in 1971 (years)
<i>a) largest decreases in training requirements</i>		
Carpenters	-5.7	6.4
Musician or composer	-5.1	6.8
Air traffic controllers	-5.0	6.5
Production supervisors or foremen	-4.7	5.4
Dental laboratory and medical appliance technicians	-4.7	5.9
Geologists	-4.5	7.5
Precision makers, repairers, and smiths	-4.4	5.9
Insurance adjusters, examiners, and investigators	-4.4	5.7
Civil engineers	-4.2	7.2
Recreation and fitness workers	-4.1	6.4
Chemical engineers	-4.0	7.0
Masons, tilers, and carpet installers	-3.9	4.7
Heating, air conditioning, and refrigeration mechanics	-3.9	5.4
Electrical engineer	-3.9	6.9
Petroleum, mining, and geological engineers	-3.8	7.1
Aerospace engineer	-3.8	6.8
Mechanical engineers	-3.8	6.8
Explosives workers	-3.8	4.4
Patternmakers and model makers	-3.7	5.2
Molders, and casting machine operators	-3.6	4.2
<i>b) largest increases in training requirements</i>		
Primary school teachers	1.2	1.8
Operations and systems researchers and analysts	1.3	4.6
Agricultural and food scientists	1.3	4.7
Archivists and curators	1.5	4.5
Managers of medicine and health occupations	1.5	6.0
Public transportation attendants and inspectors	1.9	0.1
Therapists, n.e.c.	2.3	2.9
Proofreaders	2.3	0.7
Vocational and educational counselors	2.5	4.1
Registered nurses	2.7	3.1
Social workers	2.7	3.3
Social scientists, n.e.c.	3.0	4.2
Economists, market researchers, and survey researchers	3.2	4.3
Optometrists	3.9	3.6
Pharmacists	4.3	3.2
Librarians	4.4	3.1
Podiatrists	4.5	3.0
Physical scientists, n.e.c.	4.5	3.0
Other health and therapy	4.5	3.0
Dietitians and nutritionists	4.6	2.9