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

Andy Feng & Georg Graetz †

August 17, 2013

Preliminary, comments welcome. Please do not cite without permission.

Abstract

We study the labor market effects of increased automation. We build a model in which firms optimally design machines, train workers, and assign these factors to tasks. Consistent with findings from computer science and robotics, the model features tasks which are difficult from an engineering perspective but easy for humans to carry out due to innate capacities for complex functions like vision, movement, and communication. In equilibrium, firms assign low-skill workers to such tasks. High skill workers have a comparative advantage in tasks which require much training and are difficult to automate. Workers in the middle of the skill distribution perform tasks of intermediate difficulty on both dimensions. When the cost of designing machines falls, firms adopt machines predominantly in tasks that were previously performed by middle-skill workers. Occupations at both the bottom and the top of the wage distribution experience employment gains. Wage inequality increases at the top but decreases at the bottom. As design costs fall much further, only the most skilled workers enjoy rising skill premiums, and an increasing fraction of the labor force is employed in jobs that require little or no training. The model's implications are consistent with recent evidence of job polarization and a hollowing-out of the wage distribution. In addition, we provide novel evidence on trends in occupational training requirements that is in line with the model's predictions.

*We thank Francesco Caselli, Guy Michaels, John Van Reenen, Alan Manning, Michael Boehm, Johannes Boehm, and Claudia Steinwender for valuable comments and suggestions.

†Both: London School of Economics and Center for Economic Performance, Houghton Street, London WC2A 2AE, UK. Emails: a.s.feng@lse.ac.uk and g.graetz@lse.ac.uk.

1 Introduction

How does labor-replacing technical change affect the allocation of workers to jobs, and what are its effects on the wage distribution? To answer these questions, we build a model guided by 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, pp.23-24). Second, there are tasks that seem easy to any worker but building a machine capable of performing them may be costly if not impossible: occupations such as waiters, taxi drivers, or housekeepers do not require much skill beyond vision, movement, and communication in natural language, but these are highly complex functions from an engineering point of view. The two insights combine to generate an equilibrium in which tasks that require little or no training are performed by workers at the bottom of the skill distribution—that is, workers with high learning costs. Middle skill workers compete directly with machines, as their comparative advantage (CA) is in tasks that require a non-negligible amount of training and that are of intermediate complexity in engineering terms. Finally, high skill workers' CA is in highly training-intensive, complex tasks, and thus they face significantly less competition from machines than the middle-skilled.

We model labor-replacing technical change as an exogenous fall in the cost of making machines, resulting from innovations that facilitate the automation of a wide range of tasks. Examples include the electrification of manufacturing,[1] the information and communication technology (ICT) revolution, and recent advances in robotics and artificial intelligence.[2] Responding to the fall in machine design costs, firms adopt machines in tasks that were previously performed by middle skill workers. Low skill workers' jobs might also be subject to automation, but to a lesser degree than middle skill workers. The reallocation of workers causes occupations (sets of tasks) at both the bottom and the top of the wage distribution to experience employment gains, a phenomenon known in the literature as job polarization. Wage inequality increases at the top but decreases at the bottom. As machine design costs drop further, the part of the wage distribution featuring rising inequality becomes smaller. An increasing fraction of the labor force is employed in jobs that require little or no training.

We borrow from organizational economics in modeling the production process. This allows us to work with a precise notion of objective complexity that we call *knowledge intensity*. Following Garicano (2000), we assume that production requires knowledge which must be possessed by workers and embodied in machines. The *knowledge intensity* of a task is the amount of knowledge required to attain a given level of productivity. As machines are made of inanimate matter which is devoid of knowledge to start with, it is knowledge intensity alone that determines the cost of building a machine capable of performing a task. However, the amount of training a worker requires may differ even across two tasks of equal knowledge intensity: in some cases she can draw on innate capabilities, as when driving a car safely through traffic; but in other cases knowledge must be painstakingly acquired, as when solving differential equations. The distinction between

[1] Electrification facilitated automation because electric motors could be arranged much more flexibly than steam engines (Boff 1967, p.513).

[2] We provide a list of examples for recent progress in these areas in Section 2.

knowledge intensity and *training intensity* is critical for explaining why middle skill workers are most affected by increasing automation.

Although the model assumes that all factors are perfect substitutes at the task level, it is possible for complementarities between factors to arise because tasks are q-complements in the production of the final good. When it gets cheaper to make machines, firms respond in two ways. 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. The second effect, however, forces some workers to move to different tasks, creating excess supply which puts downward pressure on their wages. Since middle skill workers are most likely to be displaced by increased automation, their wages relative to low skill and high skill workers will decline. Thus, whether technology substitutes or complements (in terms of wage effects) for a worker of given skill type will depend on that worker's exposure to automation.

Our model features a continuum of worker types as well as a continuous task space, building on the framework developed by Costinot and Vogel (2010). Existing task-based models in the wage inequality literature either assume a small number of worker types and a continuum of tasks, or a continuum of types and a small number of tasks. The disadvantage of either approach is that by construction, relative wages within large sub-groups of workers are unaffected by technical change.[3] Our assumptions allow us to characterize the effects of labor-saving innovations on the entire wage distribution, and we are able to derive predictions about changes in both between- and within-group wage inequality.[4]

The model's implications are consistent with a growing empirical literature arguing that recent technical change has led to polarization of labor markets in the US and Europe.[5] Modern ICT appears to substitute for workers in middle wage jobs, while complementing labor in high and low wage jobs, thus causing the observed reallocation of employment and the hollowing-out of the wage distribution.[6] Our model provides a precise mechanism explaining these findings. In particular, the model suggests that the ICT revolution has caused job polarization because it has facilitated a more wide-ranging automation of tasks. A corollary is that job polarization should not be a unique consequence of the recent ICT revolution. Indeed, Gray (2011) finds that electrification in the US during the first half of the 20th century led to a fall in the relative demand for middle skill workers.

Our theory delivers several novel predictions about trends in occupational training requirements. In the model we distinguish between general and specific skill. The former refers to the ease with which a worker acquires the latter, namely, task-specific knowledge. We gauge the

[3]To see this for the case of a continuum of workers and a discrete set of tasks, consider two distinct workers who are both assigned to the same task and remain so after a change in technology. The two workers' relative wage will stay constant as they both face the exact same change in the price of the task they perform.

[4]In the wage inequality literature, between-group inequality refers to differences in mean wages across groups defined by observable characteristics such as education and experience. Within-group inequality refers to wage dispersion within such groups.

[5]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 European economies by Goos, Manning, and Salomons (2009).

[6]See Autor, Levy, and Murnane (2003), Michaels, Natraj, and Van Reenen (2010), and Goos, Manning, and Salomons (2011) for evidence favoring the technological explanation.

amount of task-specific knowledge required in an occupation using measures of training intensity from the Dictionary of Occupational Titles (DOT) and the O*NET database. This allows us to measure training requirements in the US at two points in time, 1971 and 2007.

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. Furthermore, we show that occupations that initially had intermediate training intensities 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. We also find that almost all occupations experienced an increase in mean years of schooling, irrespective of changes in training requirements. This is in line with the model's prediction about an increase in skill supply. Finally, we show that changes in occupational wage premia are positively correlated with changes in training requirements, again consistent with the model.

The paper's main contributions can be summarized as follows. First, we present the first model of labor-saving technical change that allows for endogenous technology adoption as well as endogenous machine design and training choices. Second, to the best of our knowledge our model is the first to generate job polarization endogenously. Existing models[7] usually assume that technology substitutes for middle skill workers while complementing high and low skill ones—this is instead a result in our paper. Third, we provide comparative static results for the entire wage distribution, for instance we derive predictions about the effects of automation on wage inequality at the top of the wage distribution. Finally, we derive and test novel predictions about trends in occupational training requirements. The connection between technological change and training seems to have been neglected in the empirical literature (Handel 2000),[8] but our model suggests that the two topics are intimately linked.

The plan of the paper is as follows. The next subsection reviews related literature. Section 2 motivates the conceptual framework which underlies our modeling of tasks, and relates our framework to the one used by Autor, Levy, and Murnane (2003). Section 3 presents and solves the model. Section 4 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 5 presents two extensions to the model: endogenous capital accumulation and a fixed cost of technology adoption. Section 6 confronts the model's prediction with existing empirical evidence and takes novel implications of the model to the data. Section 7 concludes. All proofs are contained in the appendix.

1.1 Related literature

We build on a rather small 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

[7]See e.g. Autor, Levy, and Murnane (2003), Autor, Katz, and Kearney (2006), Acemoglu and Autor (2011), Autor and Dorn (2013), and Cortes (2012).

[8]Not so in the theoretical literature on wage inequality—see Section 1.1.

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 heterogenous 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 wider theoretical literature that has used assignment models to investigate the effects of technical change on the role of workers in the production process and on the wage distribution. One 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, following a recent literature that has emphasized a task-based approach to labor markets (Autor 2013). The interaction of workers and machines is nevertheless present in our model: since tasks are assumed to be q-complements, the efficiency of machines affects the marginal products of all workers in the economy.

We adopt the model of task production developed by Garicano (2000) in his theory of firm organization and knowledge hierarchies. Garicano and Rossi-Hansberg (2006) use this model to analyze how hierarchical organizations are affected by a decline in communication and knowledge acquisition costs, another consequence of the ICT revolution. Our focus is instead on labor-saving innovations, and we keep the model simple by not allowing hierarchies of multiple layers.

Finally, 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.

2 Motivating the Conceptual Framework

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. Steven Pinker notes that "[the] mental abilities of a four-year-old that we take for granted—recognizing a face, lifting a pencil, walking across a room, answering a question—in fact solve some of the hardest engineering problems ever conceived" (Pinker 1994, p.192). In contrast, many abilities that humans must painstakingly acquire, such as mastery in arithmetic, are trivial from an engineering perspective. This insight 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 paradox 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. While the average human will find it somewhat challenging to divide 105 by 14 in his head, he has no trouble crossing a crowded public square on foot without constantly bumping into people. However, in terms of intrinsic difficulty the latter task is much harder than the former.[9] The reason that we are usually not aware of this fact is that we rely on innate abilities[10] for functions like movement or perception, but have no such advantage when it comes to abstract tasks like arithmetic.[11]

While in reality the intrinsic difficulty of a task would have to be assessed on multiple dimensions, we adopt a one-dimensional concept for simplicity. In our framework, a task's intrinsic difficulty is measured by its *knowledge intensity*. Formally, more-knowledge-intensive tasks require a larger amount of knowledge for a given level of productivity. Solving the division exercise mentioned above is a task with low knowledge intensity, because the required procedure can easily be codified. Crossing the crowded public square, in contrast, requires a vast amount of knowledge about movement and coordination, not to mention the ability to correctly anticipate the actions of the people around.

Because machines are made of inanimate matter which is initially devoid of knowledge,[12] it is knowledge intensity alone that determines the difficulty of building a machine capable of performing a given task. The preceding discussion makes clear however that the amount of training a human worker requires may differ even across two tasks of equal knowledge intensity. This is because she can draw on a vast endowment of knowledge providing her with certain innate capabilities, although for the most part this knowledge may be unconscious or *tacit*. The presence of such knowledge endowments (either innate or acquired early) applicable to a wide range of tasks suggests introducing a second dimension into our task framework, which we call

[9] On the challenge of making walking robots, to say nothing of visual perception, Spear (2001, p.336) comments that "[in] practice this is very difficult to achieve as the leg position requires continuous sensing to ensure safe positioning and large amounts of real time computing to ensure that the robot moves without overbalancing—something the human brain achieves with ease (when sober anyway!)."

[10] "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 Mameli and Bateson (2011).

[11] 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 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."

[12] Of course, many materials have productive properties—take for instance copper with its electrical conductivity; but the 'knowledge' contained in materials is usually highly specific and limited, so that it is probably safe for our purposes to ignore this exception to the rule.

Table 1: Two-Dimensional Task Framework, Examples

		Training intensity	
		−	+
Knowledge intensity	+	driving a car language waiting tables	grading essays research strategic decision making
	−	assembly driving a train	arithmetic bookkeeping grading MCQs

training intensity: more-training-intensive tasks require more resources for equipping a human worker with a given level of knowledge specific to the task. In contrast to knowledge intensity, which refers to an objective understanding of knowledge requirements, the training intensity of a task is an attribute that only arises in the context of a worker performing a task.

We offer two illustrative examples. First, compare the task of driving a train with that of driving a car. The former takes place in a well-controlled environment, unlike the latter, which has therefore higher knowledge intensity.[13] But to humans, the two tasks may not seem all that different in terms of 'difficulty'—the uncertainties of navigating through road traffic do not pose an extraordinary challenge since many of the key functions they require, such as vision, are innate.

Second, contrast the task of grading an exam consisting of multiple choice questions (MCQs) with that of marking an essay-based test. MCQs allow only for a limited set of possible answers, and the recipe for grading them is trivial (but the task is still somewhat training intensive as it requires the ability to read and add up marks). In contrast, grading an essay may involve assessing a large variety of approaches to the questions posed. Clearly, the latter is more knowledge-intensive than the former. But in this case, it is also more training-intensive: most humans will find grading an essay the more difficult task, perhaps even impossible to complete in the absence of subject-specific training. Driverless trains and machine-grading of MCQs have been around much longer than driverless cars and automatic grading of essays, both appearing only recently (Markoff 2010, Shermis and Hamner 2012). We will show the model to be consistent with this fact. Table 1 provides an overview of our task framework and contains some further examples.

We are not the first to employ a multi-dimensional task space to analyze the impact of technological change on jobs and wages. In particular, Autor, Levy, and Murnane (2003, henceforth ALM) categorize tasks as routine and non-routine on one dimension, and as analytic, interactive and manual on another. 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

[13]We consider only the process of driving the train, not the engineering knowledge and familiarity with railway infrastructure that train drivers posses in practice.

by machines" (ibid.). The terms analytic, interactive and manual are used to characterize both routine and non-routine tasks in more detail.

While ALM's framework echoes many of the issues that we have discussed here, we believe that our own framework offers several advantages. First, it is more general, as it avoids specific attributes such as interactive and manual. Second, 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.[14] Thus, knowledge intensity is an objective, time-invariant measure of the information required to do a particular task, irrespective of whether a machine or a human does it. Third, the concept of training intensity is absent in ALM. 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.

Notwithstanding these differences, it is still possible to interpret ALM's empirical results in light of our framework. For instance, their measure of routine-ness might in practice be inversely related to knowledge intensity. We will return to this issue when discussing how our model matches up to empirical findings in Section 6.

While we believe that our task framework is an improvement over existing literature and that it generates useful and novel insights, there are some limitations. For instance, technological 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 knowledge-intensive by moving it to a more controlled environment.[15] 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 knowledge intensity. 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. There are three reasons. First, 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. Second, we can parameterize the model such that machine productivity levels in some tasks are vanishingly small. Third, and most 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 combi-

[14]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.

[15]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).

nations of the two (IBM's successful Jeopardy contestant Watson). Markoff (2012) provides an account of the increased flexibility, dexterity, and sophistication of production robots.[16] 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.

3 The Model

3.1 Overview

The model has one period which we interpret as a worker's lifetime.[17] There is a unique final good which is produced using a continuum of intermediate inputs, or *tasks*. These tasks are produced 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. Factors' productivity is not a given: intermediate firms must train workers, 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 levels and machine quality are choices faced by the intermediate firms which, unlike the decision of what factor to hire, are made independently of factor prices and task prices. This is because training and design costs are assumed to be in units of factor inputs and not in units of the final good. Optimal training and design choices, and hence productivity, result instead from the properties of tasks and their interplay with attributes of the factors of production. Characterizing these choices is subject of the Section 3.5. The result is a productivity schedule that determines comparative advantage between factors and across tasks. This then allows us to apply standard results to solve for the equilibrium assignment of factors to tasks in Section 3.6. Thus, we proceed by a kind of 'backward induction': first, we solve for factors' productivity conditional on firms' hiring these factors; and second, we characterize hiring choices, using the results of the first step.

3.2 The Task Space

Tasks differ along two dimensions, knowledge intensity, denoted by $\sigma \in \Sigma$, and training intensity, denoted by $\tau \in T$. The higher is a task's σ, the more knowledge is required for a *worker* or a *machine* to attain a given level of productivity. The higher is a task's τ, the more resources are required to equip a *worker* with a given level of knowledge. Recall that the concept of knowledge intensity refers to an objective understanding of knowledge requirements, for instance,

[16] An overview of recent developments in robotics research can be found in Nourbakhsh (2013).

[17] We discuss a dynamic (multi-period) version of the model in Section 5.1.

the amount of information processing required to perform a given task. In contrast, the training intensity of a task is an attribute that only arises in the context of a worker performing a task.

Completion of tasks results in intermediate outputs that are used to produce the final good. Let Y denote the output of the unique final good, and let task output be denoted by $y(\sigma,\tau)$. For tractability, we use a Cobb-Douglas production function,

$$\log Y = \int_{\Sigma\times T} [\log y(\sigma,\tau)]\, dB(\sigma,\tau).$$

The weighting function $B(\sigma,\tau)$ determines the relative importance of each task in final good production. To ensure constant returns to scale we assume $\int_{\Sigma\times T} dB(\sigma,\tau) = 1$.

Throughout most of our analysis we make the following, simplifying assumption about the domains of the parameters τ and σ.

Assumption 1 $\quad \tau \in T = \{0,1\}, \quad \sigma \in \Sigma = [\underline{\sigma},\overline{\sigma}], \underline{\sigma} > 0$

Under this assumption, there is a set of tasks for which $\tau = 0$, so that knowledge acquisition costs are zero, or equivalently, all workers have an innate ability to perform these tasks. We will call these tasks 'innate ability tasks'. We will refer to the tasks with $\tau = 1$ as 'training-intensive tasks'. Within both these sets of tasks, knowledge intensity varies continuously. We will state explicitly when Assumption 1 is imposed.

3.3 Worker Training, Machine Design, and Technological Change

The technologies for training workers and designing machines are as follows. Intermediate firms must pay τ/s efficiency units of labor to equip a worker of skill s with a unit measure of knowledge. Higher skilled workers have lower learning costs. Higher values of τ imply a larger learning cost, holding knowledge and skill constant.

Similarly, to transform one unit of capital into a machine equipped with a unit measure of knowledge, intermediate firms must pay $c_K \equiv 1/s_K$ units of capital. We will refer to c_K as the machine design cost, which is the main exogenous driving force in our model. As a matter of notation, it will be more convenient to work with s_K, 'machine skill', instead of c_K. Notice that a tasks's τ does not affect design costs, by definition.

Workers' and machines' productivity depends on their task-specific knowledge as well as a task-neutral productivity term, which shifts a factor's productivity proportionately in all tasks. Let task-neutral productivity of machines be denoted by A_K.

Our model admits exogenous technological change in the form of a decrease in c_K or an increase in A_K, although we will mainly be concerned with the former. A fall in c_K represents any technological advance that lowers the cost of automation of a wide range of tasks, typically a combination of improved software (programming languages, algorithms) and improved hardware (CPU speed, robotics). A rise in A_K represents improved efficiency of existing machinery. In reality, the forces affecting the two parameters may not always be mutually exclusive. This does not impair the model's ability to generate sharp predictions, however, since both parameters give rise to the same comparative statics.

3.4 A Simple Example

To illustrate how task characteristics and factor attributes affect productivity differences across factors and tasks, we present a simple example. We impose Assumption 1. Let us assume for the moment that worker training and machine design are exogenously determined by task characteristics. In particular, suppose that factors are either made capable of performing a task or not, so that there is no intensive margin for task-specific productivity. Let knowledge intensity σ be the amount of knowledge required for a factor to be able to perform a given task. A worker with learning cost $1/s$ will produce $A(1 - \sigma/s)$ units of task output in training-intensive tasks ($\tau = 1$), where A is the worker's task-neutral productivity. The same worker will produce A units in any innate ability task ($\tau = 0$). A machine will produce $A_K(1 - \sigma/s_K)$ units regardless of training intensity.

Now consider two workers with skill levels s, s' such that $s' > s$, and two tasks with equal training intensity $\tau = 1$ but different knowledge intensities σ, σ' such that $\sigma' > \sigma$. (How task-neutral productivities A and A' compare is irrelevant for what follows.) Simple algebra establishes that the higher skilled worker is relatively more productive in task σ', i.e. she has a comparative advantage in the more knowledge-intensive task. Machines' comparative advantage will depend on the level of design costs $c_K \equiv 1/s_K$. For instance, if $s_K < s$, then the machine has a comparative advantage over both workers in the less knowledge-intensive task.

Next, take an innate ability task and a training-intensive task both with equal knowledge intensity σ. Machines are equally productive in both tasks but workers are more productive in the innate ability task. Therefore, machines have a comparative advantage in the training-intensive tasks. This is why some training-intensive tasks will always be performed by machines, even if machine design cost exceed the training cost of the least-skilled worker.

Finally, consider again two workers with skill levels s, s' such that $s' > s$ and take an innate ability task and a training-intensive task both with equal knowledge intensity σ. Because the higher-skilled worker has a higher *task-specific* productivity in the training-intensive task, she has a comparative advantage in that task. This is why workers at the bottom of the skill distribution will generally perform innate ability tasks, and why middle skill workers will compete with machines in training-intensive tasks of intermediate knowledge intensity.

The simple example illustrates the main forces driving our results about the effects of increased automation on job assignment and the wage distribution. In fact, the simple model presented here generates an equilibrium assignment and comparative static results that are qualitatively the same as in the model with endogenous worker training and machine design. However, the simple model does not explicitly describe the production process, so that it is not clear what precisely drives the results. Moreover, it does not allow us to assess if the results are robust to allowing firms a productivity choice (via worker training and machine design). We address these limitations in the following section.

3.5 The Production Process for Tasks and Firms' Productivity Choices

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 3.3 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 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 knowledge intensity σ is given by the cumulative density function $F(Z; \sigma)$, which we assume to be continuously differentiable in both Z and the shift parameter σ. Let $\partial F/\partial\sigma < 0$, so that σ indexes first-order stochastic dominance. In terms of the examples discussed in Section 3.2, driving a car and grading an essay are more knowledge-intensive (higher σ) 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. Let $\varepsilon_{F,\sigma}(Z, \sigma)$ denote the elasticity of F with respect to σ holding Z constant, and similarly for $\varepsilon_{f,\sigma}(Z, \sigma)$. We impose the following condition on the family of distributions $F(Z; \sigma)$.

Assumption 2 $\varepsilon_{F,\sigma}(Z, \sigma) < \varepsilon_{f,\sigma}(Z, \sigma)$ *for all* $Z, \sigma > 0$

This assumption will give rise to a set of intuitive comparative advantage properties, for instance high skill workers will have a comparative advantage in knowledge-intensive tasks. One of the distributions satisfying Assumption 2 is the exponential distribution with mean σ.

Note that the distribution of problems depends only on σ and not on τ. 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) knowledge intensity.

We now characterize optimal training an design choices and thus 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. Assume that each worker is endowed with one efficiency unit of labor. After incurring learning costs, $1 - \tau z/s$ 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 (σ, τ) be denoted by $\alpha^N(s, \sigma, \tau)$, 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

$$\begin{aligned} \alpha^N(s,\sigma,\tau) &\equiv \sup_z F(z;\sigma)\left[1-\tfrac{\tau}{s}z\right], \\ \alpha^K(s_K,\sigma) &\equiv \sup_z F(z;\sigma)\left[1-\tfrac{1}{s_K}z\right], \end{aligned}$$

A unique interior solution to the worker training problem exists provided $\tau > 0$, while the machine design problem always admits a unique interior solution.[18] The optimal knowledge levels $z^N(s,\sigma,\tau)$ and $z^K(s_K,\sigma)$ are pinned down by the first-order conditions

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

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.

We will formalize the concept of innateness by assuming that some tasks feature $\tau = 0$. It is immediate that in such innate ability tasks, $\alpha^N(s,\sigma,0) = 1$. Thus, optimal worker and machine productivities are given by

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

and

$$\alpha^K(s_K,\sigma) = F(z(s_K,\sigma);\sigma)\left[1-\frac{1}{s_K}z(s_K,\sigma)\right].$$

We impose Assumption 1 for the remainder of the paper. Let the set of worker skills be given by $S = [\underline{s},\overline{s}]$ and let $\breve{s}$ be an element in set $\breve{S} = s_K \cup S$. By the above equations, we have that $\alpha^N(\breve{s},\sigma,1) \equiv \alpha^K(\breve{s},\sigma)$. Thus, workers and machines face the same productivity schedule in training-intensive tasks. We drop superscripts and define the function

$$\alpha(\breve{s},\sigma) = F(z(\breve{s},\sigma);\sigma)\left[1-\frac{1}{\breve{s}}z(\breve{s},\sigma)\right] \qquad \breve{s} \in \breve{S} = s_K \cup [\underline{s},\overline{s}], \tag{2}$$

where $z(\breve{s},\sigma)$ is implicitly given by (1) when $\tau = 1$.

We now turn to the properties of the productivity schedule $\alpha(\breve{s},\sigma)$. First notice that $\alpha \in$

[18] 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.

$(0, 1)$ by (2). Furthermore, from applying the envelope theorem to (2) it follows that α is increasing in $\breve{s}$ and decreasing in σ. Higher skilled factors are more productive since they face a lower learning/design cost, and productivity declines in knowledge intensity since a larger cost is incurred to achieve a given level of productivity. To characterize comparative advantage, we rely on the following result.

Lemma 1 *The productivity schedule $\alpha(\breve{s}, \sigma)$ is strictly log-supermodular if Assumption 2 holds.*

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

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

For instance, high skill workers have a comparative advantage over low skill workers in more knowledge-intensive tasks; all workers with $s > s_K$ have a comparative advantage over machines in more knowledge-intensive tasks; and so on. As the proof of Lemma 1 establishes, these comparative advantage properties hold if and only if optimal knowledge $z(\breve{s}, \sigma)$ is increasing in σ. Thus, high skill factors have a comparative advantage in more knowledge-intensive tasks because these tasks induce a higher level of knowledge, and to high skill factors this comes at a lower cost.

The effect of σ on the optimal knowledge level is in principle ambiguous. A higher σ implies a lower opportunity cost of learning an additional problem since factors are less productive, ceteris paribus. However, the marginal benefit may increase or decrease depending on the problem distribution. Assumption 2 ensures that the fall in marginal costs outweighs any effect on the marginal benefit.

Comparative advantage properties regarding training intensity are straightforward. Since α is increasing in $\breve{s}$, 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 knowledge-intensity 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 knowledge intensity. 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 4 below.

3.6 Competitive Equilibrium

To complete the setup of the model, 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 differentiable CDF, and $v(s)$ the PDF, both with support $S = [\underline{s}, \overline{s}]$. Let the share of innate ability tasks ($\tau = 0$) in final good production be β. The production function can now be written as

$$\log Y = \frac{1}{\mu} \int_{\underline{\sigma}}^{\overline{\sigma}} \{\beta \log y_0(\sigma) + (1-\beta) \log y_1(\sigma)\} \, d\sigma, \tag{3}$$

where the term $\mu \equiv \overline{\sigma} - \underline{\sigma}$ ensures constant returns to scale. The subscripts 0 and 1 indicate innate ability ($\tau = 0$) and training-intensive ($\tau = 1$) tasks, respectively.

We have established in Section 3.5 that in innate ability tasks, machine productivity is given by $\alpha(s_K, \sigma)$, while worker productivity equals one. Hence, output of the innate ability task with knowledge intensity σ is given by

$$y_0(\sigma) = A_K \alpha(s_K, \sigma) k_0(\sigma) + \int_{\underline{s}}^{\overline{s}} n_0(s, \sigma) d\sigma, \tag{4}$$

where $k_0(\sigma)$ and $n_0(c, \sigma)$ are the masses of machine capital and of worker type s, respectively, allocated to innate ability task σ. In training-intensive tasks, as we have seen, both machine and worker productivity depends on the function $\alpha(\check{s}, \sigma)$. Hence we can write task output of the training-intensive task σ as

$$y_1(\sigma) = A_K \alpha(s_K, \sigma) k_1(\sigma) + \int_{\underline{s}}^{\overline{s}} \alpha(s, \sigma) n_1(s, \sigma) d\sigma. \tag{5}$$

There is a large number of perfectly competitive firms producing the final good, and buying task output from perfectly competitive intermediates producers. We normalize the price of the final good to one and denote the price of task σ in 'sector' $\tau \in \{0, 1\}$ by $p_\tau(\sigma)$. Profits of final good firms are given by

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

and profits of intermediate producers in sector j and with knowledge intensity σ are

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

where r is the rental rate of capital and $w(s)$ is the wage paid to a worker with skill s. Recall that design and learning costs are already included in the $\alpha(\check{s}, \sigma)$ terms which enter intermediate producer's profits through the task production functions (4) and (5).

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

$$y_0(\sigma) = \frac{\beta}{\mu} \frac{Y}{p_0(\sigma)}, \qquad y_1(\sigma) = \frac{1-\beta}{\mu} \frac{Y}{p_1(\sigma)}. \tag{6}$$

Profit maximization by intermediates producers implies

$$\begin{array}{rcll} p_0(\sigma) & \leq & w(s) & \forall s \in [\underline{s}, \overline{s}], \\ p_1(\sigma)\alpha(s,\sigma) & \leq & w(s) & \forall s \in [\underline{s}, \overline{s}], \\ p_\tau(\sigma)\alpha(s_K,\sigma) & \leq & r/A_K & \forall \tau \in \{0,1\}; \\ p_0(\sigma) & = & w(s) & \text{if } n_0(s,\sigma) > 0, \\ p_1(\sigma)\alpha(s,\sigma) & = & w(s) & \text{if } n_1(s,\sigma) > 0, \\ p_\tau(\sigma)\alpha(s_K,\sigma) & = & r/A_K & \text{if } k_\tau(\sigma) > 0. \end{array} \tag{7}$$

Factor market clearing conditions are

$$v(s) = \sum_\tau \int_{\underline{\sigma}}^{\overline{\sigma}} n_\tau(s,\sigma)d\sigma \quad \text{for all } s \in [\underline{s}, \overline{s}] \tag{8}$$

and

$$K = \sum_\tau \int_{\underline{\sigma}}^{\overline{\sigma}} k_\tau(\sigma)d\sigma. \tag{9}$$

A *competitive equilibrium* in this economy is a set of functions $y : \Sigma \times T \to \mathbb{R}^+$ (task output); $k : \Sigma \times T \to \mathbb{R}^+$ and $n : S \times \Sigma \times T \to \mathbb{R}^+$ (factor assignment); $p : \Sigma \times T \to \mathbb{R}^+$ (task prices); $w : S \to \mathbb{R}^+$ (wages); and a real number r (rental rate of capital) such that conditions (1), (2), and (4) to (9) hold.

The equilibrium assignment of factors to tasks is determined by comparative advantage, which is a consequence of the zero-profit condition (7).[19] Because high skill workers have a comparative advantage in training-intensive tasks (holding knowledge intensity 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.

We focus on the empirically relevant case in which machines as well as workers perform both training-intensive and innate ability tasks.[20] In this case, machines are assigned to a subset of innate ability and training-intensive tasks that are relatively less knowledge-intensive, while low skill workers perform the remaining innate ability tasks: there is a threshold task σ_0^*, the marginal innate ability tasks, dividing the set of innate ability tasks into those performed by

[19]To see how comparative advantage determines patterns of specialization, consider two firms, one producing training-intensive task σ, the other producing training-intensive task σ'. Suppose in equilibrium, firm σ is matched with workers of type s and firm σ' is matched with workers of type s'. Then (7) 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 σ (σ'), precisely the task to which she was assumed to be matched.

[20]Sufficient conditions for the existence of such an equilibrium are derived Appendix A.1. We assume throughout that these conditions are satisfied. We note however that in general, no innate ability tasks may be performed by machines, and/or no training-intensive tasks may be performed by workers.

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 σ_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 3.5. These properties also imply $\sigma_0^* < \sigma_1^*$: the marginal training-intensive task is always more knowledge-intensive than the marginal innate ability task (recall that machines are relatively more productive in training-intensive tasks than workers, holding knowledge intensity 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* $s^* \in (\underline{s}, \overline{s}]$ *such that*

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

(b) *If* $k_0(\sigma) > 0$ *for some* σ*, then* $s^* > s_K$*, and there exist* $\sigma_0^*, \sigma_1^* \in \Sigma$ *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 knowledge-intensive tasks. Formally, we have:

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

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 σ_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

$$M'(s) = \frac{\mu}{1-\beta} \frac{w(s)v(s)}{Y}, \tag{10}$$

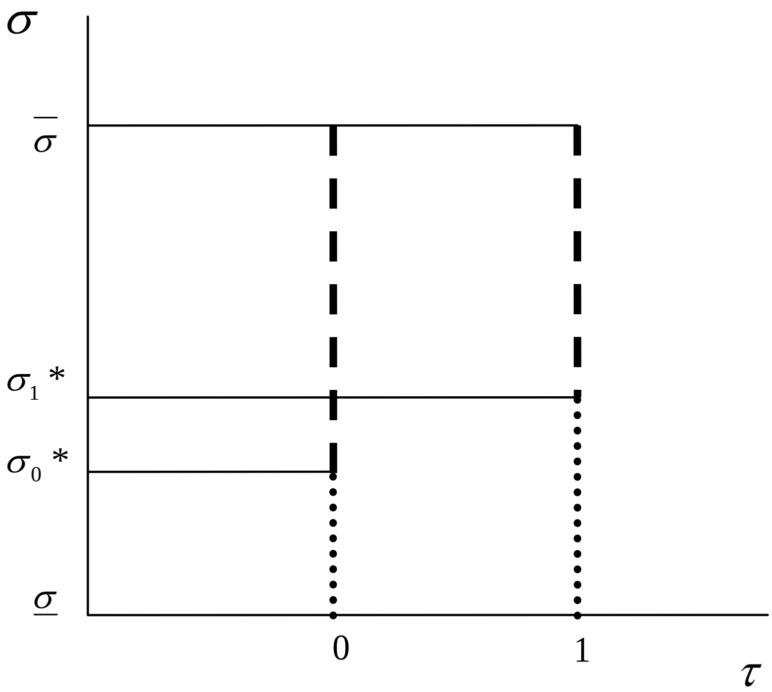


Figure 1: Assignment of labor and capital to tasks. Knowledge intensity σ is plotted on the vertical axis, while training intensity τ is plotted on the horizontal axis. Dotted and dashed lines indicate tasks performed by machines and workers, respectively.

and that the wage schedule is given by

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

The last equation is due to the fact that in equilibrium, a firm producing training-intensive task σ 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 (8) given Lemma 2, and using (6) and (7). In particular, Lemma 2 and (8) 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 (5) we obtain

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

After eliminating task output and price using (6) and (7), (10) follows. 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 σ_0^*, σ_1^*, and s^*. These equations are due to a set of no-arbitrage conditions. In particular, firms producing the marginal 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 (7) 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

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

and

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

and the no-arbitrage condition for the marginal worker is

$$w(s)=w(s^*)\quad \textit{for all } s\le s^*. \tag{15}$$

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. Relaxing this assumption would complicate the analysis, although the main results would go through as long as A is constant across tasks for each worker.

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 (but not across the two sectors unless $\beta = 0.5$). Some algebra shows[21] that machines produce task outputs

$$\begin{aligned} y_0(\sigma) &= \frac{\beta A_K \alpha(s_K, \sigma) K}{\beta(\sigma_0^* - \underline{\sigma}) + (1-\beta)(\sigma_1^* - \underline{\sigma})} \quad \text{for all } \sigma \in [\underline{\sigma}, \sigma_0^*], \\ y_1(\sigma) &= \frac{(1-\beta) A_K \alpha(s_K, \sigma) K}{\beta(\sigma_0^* - \underline{\sigma}) + (1-\beta)(\sigma_1^* - \underline{\sigma})} \quad \text{for all } \sigma \in [\underline{\sigma}, \sigma_0^*]. \end{aligned} \tag{16}$$

Using these equations to solve for the task prices in (6), and plugging the obtained expression into (7), yields

$$r = \frac{\beta(\sigma_0^* - \underline{\sigma}) + (1-\beta)(\sigma_1^* - \underline{\sigma})}{\mu} \times \frac{Y}{K}. \tag{17}$$

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

[21] By (6) and (7), we have

$$\frac{y_\tau(\sigma)}{y_\tau(\sigma')} = \frac{\alpha(s_K, \sigma)}{\alpha(s_K, \sigma')}, \qquad \frac{y_0(\widetilde{\sigma})}{y_1(\widetilde{\sigma}')} = \frac{\beta}{1-\beta} \frac{\alpha(s_K, \widetilde{\sigma})}{\alpha(s_K, \widetilde{\sigma}')}$$

for any tasks $(\sigma, \sigma', \widetilde{\sigma}, \widetilde{\sigma}')$ performed by machines. But (4), (5), and Lemma 2 imply

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

The previous two equations together give $k_\tau(\sigma) = k_\tau(\sigma')$ and $k_0(\widetilde{\sigma}) = \frac{\beta}{1-\beta} k_1(\widetilde{\sigma}')$. By (9) and Lemma 2,

$$k_0(\sigma) = \frac{\beta K}{\beta(\sigma_0^* - \underline{\sigma}) + (1-\beta)(\sigma_1^* - \underline{\sigma})} \quad \text{for all } \sigma \in [\underline{\sigma}, \sigma_0^*]$$

and

$$k_1(\sigma) = \frac{(1-\beta) K}{\beta(\sigma_0^* - \underline{\sigma}) + (1-\beta)(\sigma_1^* - \underline{\sigma})} \quad \text{for all } \sigma \in [\underline{\sigma}, \sigma_1^*].$$

market clearing conditions (4) and (8),[22]

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

Proceeding as above when solving for r, we obtain

$$w(s^*) = \frac{\beta(\overline{\sigma} - \sigma_0^*)}{\mu} \times \frac{Y}{V(s^*)}. \tag{19}$$

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

$$\frac{A_K \alpha(s_K, \sigma_0^*) K}{\beta(\sigma_0^* - \underline{\sigma}) + (1-\beta)(\sigma_1^* - \underline{\sigma})} = \frac{V(s^*)}{\beta(\overline{\sigma} - \sigma_0^*)}. \tag{20}$$

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

$$\alpha(s_K, \sigma_1^*) = \alpha(s_K, \sigma_0^*)\alpha(s^*, \sigma_1^*). \tag{21}$$

Lastly, (10) and (19) imply

$$M'(s^*) = \frac{\beta(\overline{\sigma} - \sigma_0^*)}{1-\beta} \frac{v(s^*)}{V(s^*)}. \tag{22}$$

Equations (3), (10), (11), (20), (21), and (22) together with the boundary conditions $M(s^*) = \sigma_1^*$ and $M(\overline{s}) = \overline{\sigma}$, uniquely pin down the equilibrium objects σ_0^*, σ_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 both employed in training-intensive tasks,

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

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,

$$Ew = \frac{\beta(\overline{\sigma} - \sigma_0^*) + (1-\beta)(\overline{\sigma} - \sigma_1^*)}{\mu} \times Y. \tag{24}$$

[22]Under Lemma 2, integrating (8) yields

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

but using (4) and the fact that task output is a constant y_0 results in

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

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.

4 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.

4.1 Technical Change

Consider a fall in the machine design cost from c_K to $\widehat{c}_K$, so that $\widehat{s}_K > s_K$. Let M and $\widehat{M}$ be the corresponding matching functions, and similarly for σ_0^* and $\widehat{\sigma}_0^*$; σ_1^* and $\widehat{\sigma}_1^*$; and s^* and $\widehat{s}^*$. We now state the main result of the paper.

Proposition 1 *Suppose $\widehat{c}_K < c_K$ and so $\widehat{s}_K > s_K$. Then $\widehat{\sigma}_1^* > \sigma_1^*$ and $\widehat{M}(s) > M(s)$ for all $s \in [\max\{s^*, \widehat{s}^*\}, \overline{s})$. If $\widehat{s}_K \geq s^*$, then $\widehat{s}^* > s^*$.*

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 $\sigma_0^* < \sigma_1^*$.[23] 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 σ_0^* is ambiguous.

We are unable to rule out the possibility that a very small decrease in the machine design cost leads to the marginal worker becoming less skilled, $\widehat{s}^* < s^*$. However, if the fall in the machine design cost is large enough, then the marginal worker becomes more skilled. Since s^* is a continuous function of c_K, a positive relationship between c_K and s^* should be restricted to a small subset of the parameter space. Thus, we limit our attention to the case where a fall in c_K triggers a rise in s^*. This implies a reassignment of some workers to innate ability tasks. Importantly, workers remaining in training-intensive also experience displacement, as they are reassigned to tasks of higher knowledge intensity due to the upward shift of the matching function. In sum, employment in tasks previously performed by low skill workers $s \leq s^*$ increases; employment in tasks previously carried out by middle skill workers $s \in (s^*, \widehat{s}^*)$ decreases; and employment in tasks formerly performed by high skill workers $s \geq \widehat{s}^*$ increases. Thus, a fall in the machine design cost causes job polarization. These effects are illustrated by Figure 2.

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

[23]Because $\sigma_0^* < \sigma_0^*$ and due to the log-supermodularity of α, the ratio $\alpha(s_K, \sigma_1^*)/\alpha(s_K, \sigma_0^*)$ is increasing in s_K.

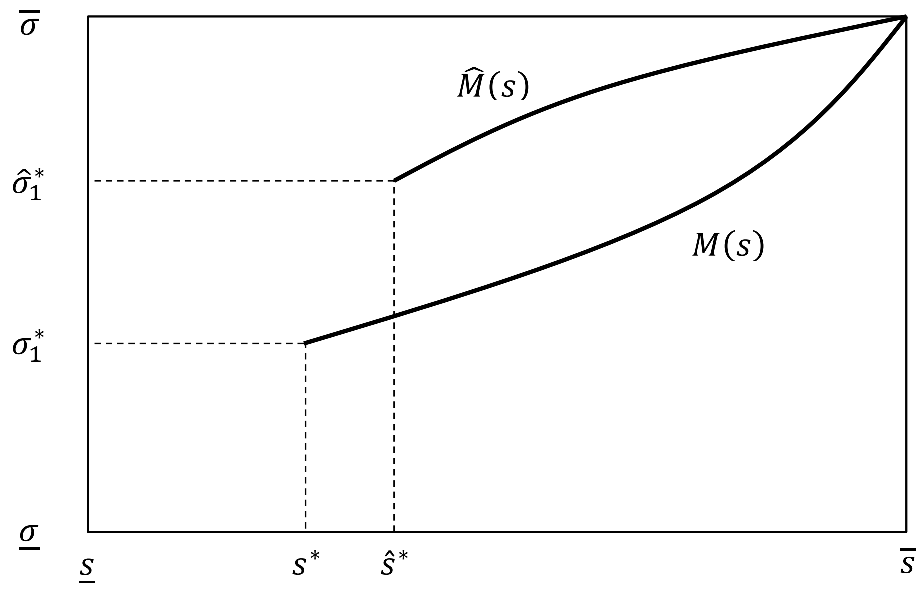


Figure 2: Assignment of workers to training-intensive tasks and the effects of technical change. Knowledge intensity σ is plotted on the vertical axis, while skill level s is plotted on the horizontal axis. The upward shift of the matching function and the shift of its lower end to the northeast are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

wages. Intuitively, since 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 they would have worked for their new employers even before the shift. 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. Thus, wage inequality rises at the top, but falls at the bottom of the distribution. This is illustrated by Figure 3. The formal result is as follows.

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

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

and

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

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 $\widehat{c}_K < c_K$. The labor share decreases.*

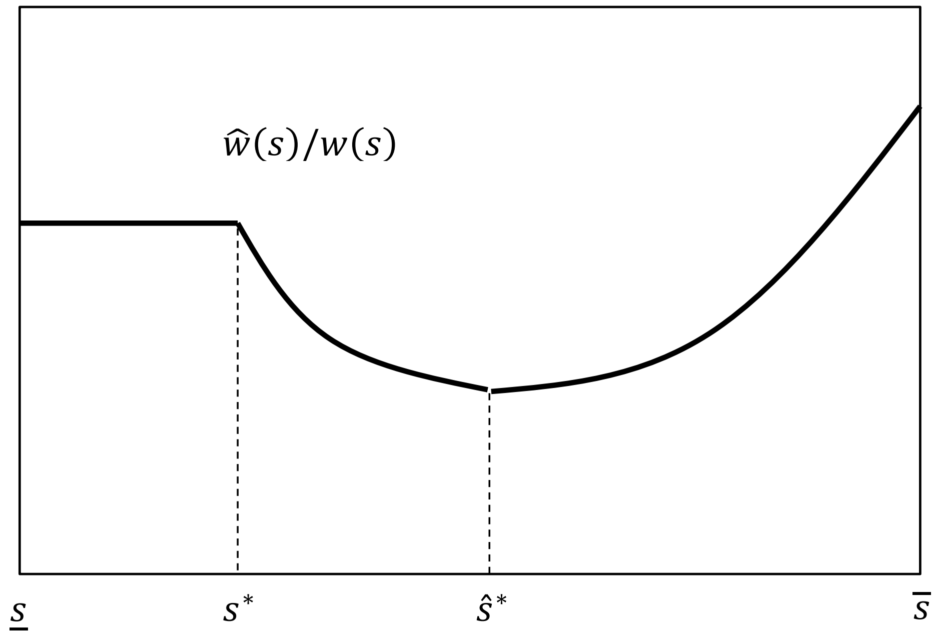


Figure 3: Changes in wages as a result of a fall in the machine design cost from c_K to $\widehat{c}_K$. For each skill level s, the ratio of new to old wages is plotted. Workers with $s \in [\widehat{s}^*, \overline{s}]$ remain in training-intensive tasks and experience a rise in the skill premium.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

4.2 Increase in Skill Abundance

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

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

For simplicity, we restrict attention to distributions with common support, and we assume that $\widehat{v}(\overline{s}) > v(\overline{s})$. 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 $\widehat{V} \succeq V$ and $\widehat{v}(\overline{s}) > v(\overline{s})$. If this change in skill endowments induces an increase in the share of income accruing to labor, then $\widehat{\sigma}_1^* < \sigma_1^*$, $\widehat{s}^* > s^*$ and $\widehat{M}(s) < M(s)$ for all $s \in [\widehat{s}^*, \overline{s})$.*

Intuitively, such a change to the distribution of skills should raise the labor share, because the labor share in our model 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.[24]

[24]The labor share is given by $\int_{\underline{s}}^{\overline{s}} \frac{w(s)}{Y} dV(s)$. Because $\widehat{V}$ first-order stochastically dominates V and $w(s)/Y$ is an increasing function, we have $\int_{\underline{s}}^{\overline{s}} \frac{w(s)}{Y} d\widehat{V}(s) > \int_{\underline{s}}^{\overline{s}} \frac{w(s)}{Y} dV(s)$. Thus, for the labor share to *decrease*, there would

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 for all s, s' with $s' > s \geq s^*$,*

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

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

5 Extensions

5.1 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[25] 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.[26] 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

need to be a sufficiently large decline in wage-output ratios for a subset of workers.

[25] Alternatively, we could assume that the economy is open to world capital markets, where it is a price taker.

[26] 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.

increase in wage inequality will be more pronounced in the long run as a result of the higher capital stock.

5.2 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.[27] 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 B 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 (knowledge intensity) of the task, 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 knowledge intensity. 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-knowledge-intensive tasks. This incentive is stronger in training-intensive tasks: as knowledge-intensity 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 are currently working on a numerical solution to verify the intuition.

6 Empirical Support for the Model's Predictions

Section 4.1 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.

[27] 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.

6.1 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 knowledge intensity, 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 knowledge intensity (though not necessarily low training intensity). The measured shift away from routine tasks is then consistent with our prediction of a reallocation towards more-knowledge-intensive 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 (2010) 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 high training- and knowledge-intensive tasks.

Wages.—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 assigned to training-intensive tasks.

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 and Dorn (2013) document that in the US over the past three decades, 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 along the lines of our model.

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.

Historical evidence.—Gray (2011) shows that electrification in the US during the first half of the 20th century led to a fall in relative demand for tasks performed by middle skill workers, providing support for the model's prediction that job polarization is not a unique consequence of the IT 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-knowledge-intensive, 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.

6.2 Trends in Occupational Training Requirements

In the model, training levels (knowledge) vary systematically with task characteristics. In particular, tasks with higher knowledge intensity require more training in equilibrium, provided $\tau > 0$.

And holding knowledge intensity constant, tasks with lower training intensity induce a lower training investment. In the extreme case of our innate ability tasks, the training investment is zero.

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 knowledge-intensive, 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.[28] 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 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 C for details. Tables C.2 and C.3 list the twenty most and least training intensive occupations in 1971 and 2007, respectively.

The definition of SVP matches our concept of task-specific knowledge 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 knowledge intensity 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

[28] 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.

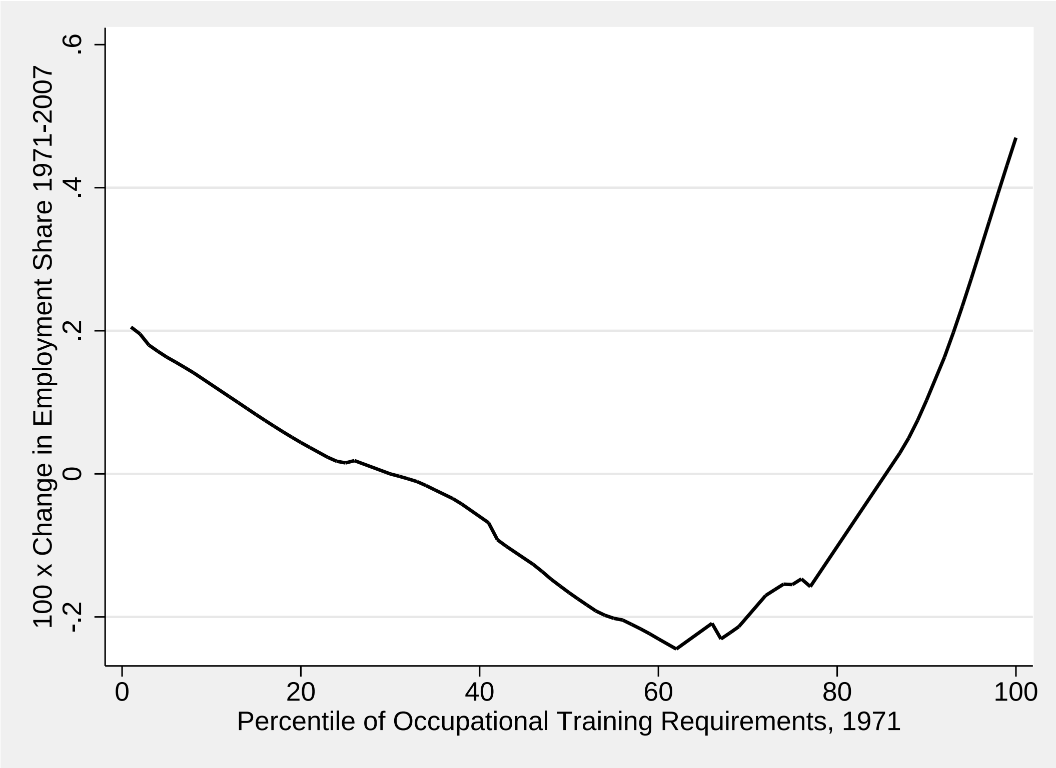


Figure 4: Changes in occupational shares, where occupations are ordered by percentile rank of the average 1980 occupational SVP-score.

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.[29] The pattern is consistent with the model's prediction of polarization of training requirements.

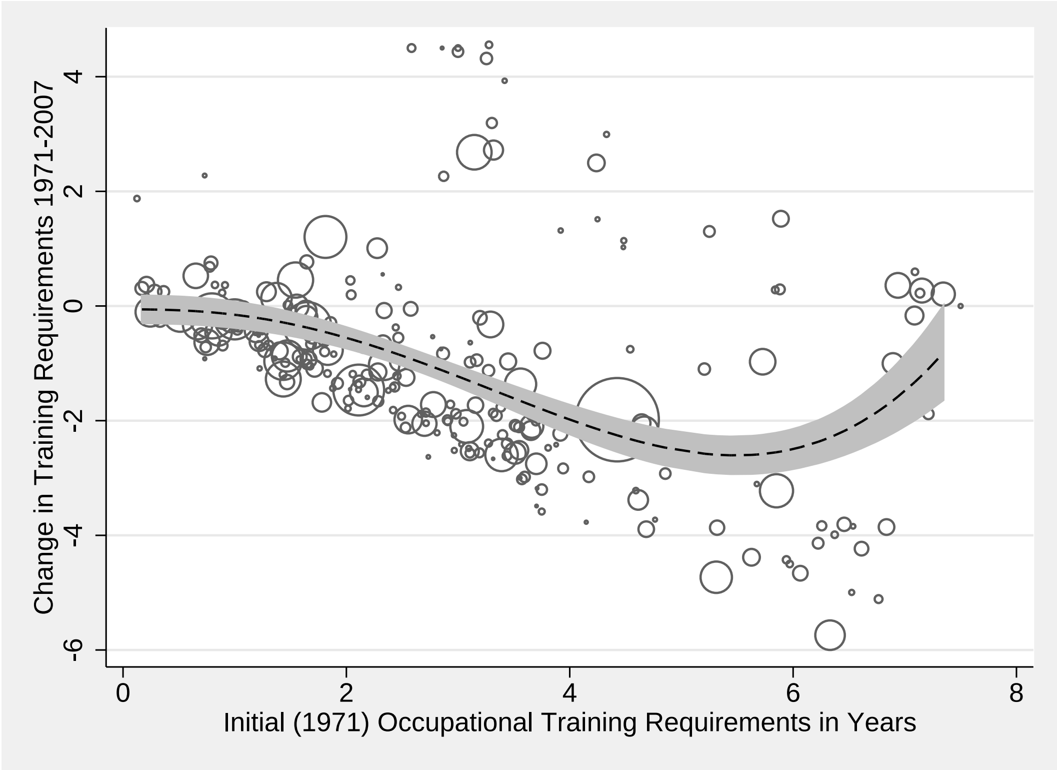
Second, the model can potentially help to make sense of changes in training requirements within occupations. If an occupation consists of a large fraction of tasks with intermediate knowledge intensity, then we would expect training requirements to decrease as these tasks are automated. Panel a) of Figure 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 C.4), 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.

Finally, 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

[29]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

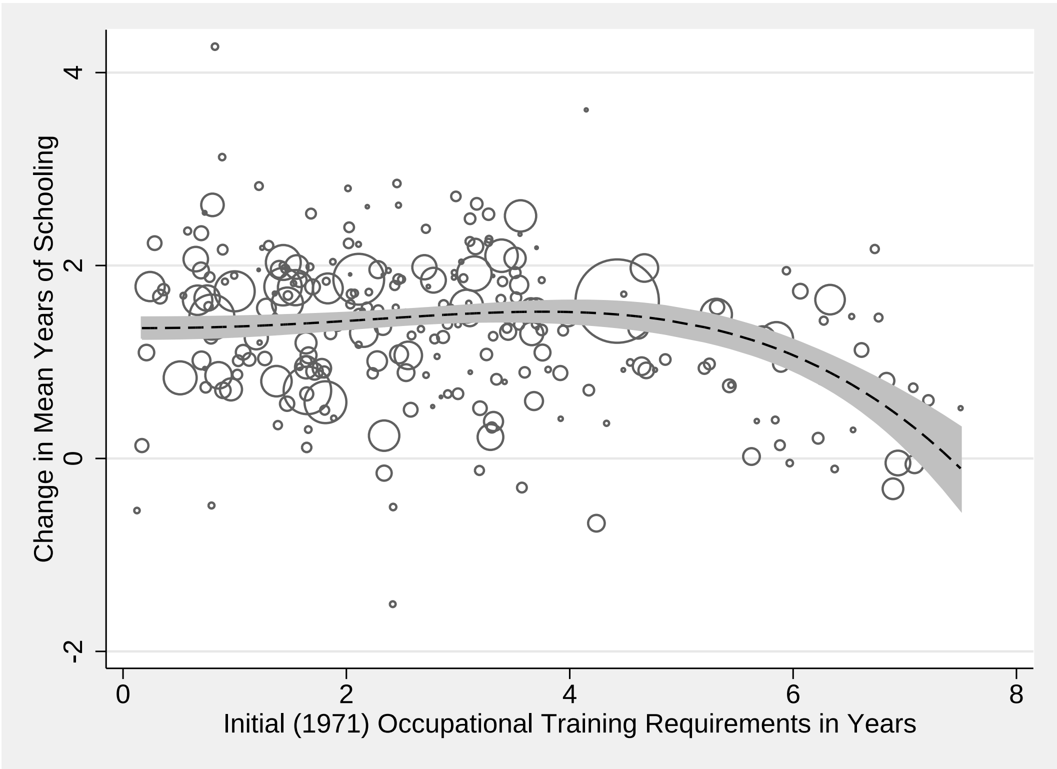


Figure 5: Changes in occupational training requirements and average years of schooling. Training requirements are calculated based on the variable *specific vocational preparation* (SVP) from the Dictionary of Occupational Titles and the O*NET database. Observations are weighted by average occupational employment shares. Fitted curves are fractional polynomials, drawn using Stata's *fpfitci* option.

Figure 6: 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 standard error of 0.026.

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). Raw data and fitted line are plotted in Figure 6. Including changes in log year 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 increased automation of tasks. For concreteness, consider an occupation whose task bundle initially includes training-intensive tasks with knowledge intensities between σ_1^* and $\sigma' > \widehat{\sigma}_1^*$. Let s' be the skill level of the worker initially performing task σ'. 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.

7 Conclusion

In this paper we make four main contributions. First, we present a model of labor-saving technical change that endogenizes firms' decisions about what tasks to automate, as well as choices of machine design and worker training. Second, we generate job polarization endogenously. We show that job polarization and a hollowing out of the wage distribution result from any technological advance that facilitates automating a broad range of tasks, and is thus not specific to the recent information technology revolution. Third, our model allows us to investigate the effects of job polarization on wage inequality near the top of the distribution, and it generates predictions about how high skill workers might be affected by further advances in AI and robotics. Fourth, the model predicts changes in occupational training requirements that are consistent with novel evidence we present. Our model does not allow for changes in the economy's task mix or changes in firm organization resulting from technological change—further research is necessary to determine whether our results are robust to these extensions.

References

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.

Autor, D. (2013): "The Task Approach to Labor Markets: An Overview," IZA Discussion Papers 7178, Institute for the Study of Labor (IZA).

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): "Has Job Polarization Squeezed the Middle Class? 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.

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.

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.

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. (2011): "Taking Technology to Task: The Skill Content of Technological Change in Early Twentieth Century United States," Working Papers 0009, European Historical Economics Society (EHES).

Handel, M. J. (2000): "Trends in Direct Measures of Job Skill Requirements," Economics Working Paper Archive 301, Levy Economics Institute, The.

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.

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. (2010): "Google Cars Drive Themselves, in Traffic," *The New York Times*, Oct 9.

Markoff, J. (2012): "Skilled Work, Without the Worker," *The New York Times*, Aug 18.

Michaels, G., A. Natraj, and J. Van Reenen (2010): "Has ICT Polarized Skill Demand? Evidence from Eleven Countries over 25 Years," NBER Working Papers 16138, National Bureau of Economic Research, Inc.

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," .

Nourbakhsh, I. R. (2013): *Robot Futures*. MIT Press.

Pinker, S. (1994): *The Language Instinct: the New Science of Language and Mind*. Penguin.

Sampson, T. (2012): "Selection into Trade and Wage Inequality," CEP Discussion Papers dp1152, Centre for Economic Performance, LSE.

Shermis, M. D., and B. Hamner (2012): "Contrasting State-of-the-Art Automated Scoring of Essays: Analysis," Discussion paper.

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.

Spear, B. (2001): "Robots and patents," *World Patent Information*, 23(4), 333–338.

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 Proofs of Formal Results Stated in the Text

A.1 Sufficient Conditions for Existence of an Interior Equilibrium

We derive sufficient conditions ensuring that an interior equilibrium with $\sigma_0^*, \sigma_1^* \in (\underline{\sigma}, \overline{\sigma})$ and hence $s^* \in (\underline{s}, \overline{s})$ 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 $(\overline{s}, \underline{\sigma}, \overline{\sigma})$.

In any equilibrium in which $k_0(\sigma) = 0$ for all $\sigma \in [\underline{\sigma}, \overline{\sigma}]$, we have by (7)

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

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

$$\alpha(s_K, \underline{\sigma}) \leq \frac{(1-\beta)(\sigma_1^* - \underline{\sigma})}{\beta(\overline{\sigma} - \underline{\sigma})} \times \frac{V(s^*)}{A_K K}.$$

The RHS of the last inequality is strictly less than $(1-\beta)/(\beta A_K K)$, hence a sufficient condition to rule out any equilibrium in which $k_0(\sigma) = 0$ for all $\sigma \in [\underline{\sigma}, \overline{\sigma}]$ is $\alpha(s_K, \underline{\sigma}) > (1-\beta)/(\beta A_K K)$ or

$$A_K K > \frac{1-\beta}{\beta} \frac{1}{\alpha(s_K, \underline{\sigma})}. \tag{25}$$

And in any equilibrium in which $n_1(s, \sigma) = 0$ for all $s \in [\underline{s}, \overline{s}]$ and $\sigma \in [\underline{\sigma}, \overline{\sigma}]$ we have by (7)

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

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

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

The RHS of the last inequality is strictly greater than $(1-\beta)/(\beta A_K K)$, hence a sufficient condition to rule out any equilibrium in which $n_1(s, \sigma) = 0$ for all $s \in [\underline{s}, \overline{s}]$ and $\sigma \in [\underline{\sigma}, \overline{\sigma}]$ is $\alpha(s_K, \overline{\sigma})/\alpha(\overline{s}, \overline{\sigma}) < (1-\beta)/(\beta A_K K)$ or

$$A_K K < \frac{1-\beta}{\beta} \frac{\alpha(\overline{s}, \overline{\sigma})}{\alpha(s_K, \overline{\sigma})}. \tag{26}$$

Combining (25) and (26), we conclude that if

$$A_K K \in S, \quad S \equiv \frac{1-\beta}{\beta}\left(\frac{1}{\alpha(s_K, \underline{\sigma})}, \frac{\alpha(\overline{s}, \overline{\sigma})}{\alpha(s_K, \overline{\sigma})}\right),$$

then the equilibrium is interior with $\sigma_0^*, \sigma_1^* \in (\underline{\sigma}, \overline{\sigma})$ and hence $s^* \in (\underline{s}, \overline{s})$. Existence of an interior equilibrium is ensured by choosing parameter values for $(\overline{s}, \underline{\sigma}, \overline{\sigma})$ such that S is a non-empty set. Our claim that the restrictions on $A_K K$ are mild given a particular choice of $(\overline{s}, \underline{\sigma}, \overline{\sigma})$ is justified if we assume that $\underline{\sigma}$ is sufficiently small so that $F(Z; \underline{\sigma})$ is close to one even for very small Z; and that $\overline{\sigma}$ is sufficiently large so that $F(Z; \overline{\sigma})$ is close to zero even for very large Z, while at the same time $\overline{s}$ is sufficiently large so that $\alpha(\overline{s},\overline{\sigma})$ stays finite. If so, then $S \to \frac{1-\beta}{\beta}(1,\infty)$.

A.2 Proofs of Lemmas Stated in the Text

Proof of Lemma 1 The productivity schedule α is strictly log-supermodular if and only if

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

Applying the envelope theorem to (2) yields

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

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

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

Thus, α is log-supermodular if and only if optimal knowledge levels are increasing in σ. Differentiating the FOC (1) yields

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

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

$$\frac{\partial}{\partial \sigma} z(\breve{s},\sigma) > 0 \quad \Leftarrow \quad \varepsilon_{F,\sigma} < \varepsilon_{f,\sigma} \text{ for all } Z, \sigma > 0. \quad \blacksquare$$

Proof of Lemma 2 (a) For any vectors (s,σ) and (s',σ') such that $n_0(s,\sigma) > 0$ and $n_1(s',\sigma') > 0$ we have by the zero-profit condition (7) $p_0(\sigma) = w(s)$ 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 α is increasing in s we must have $s' \geq s$. Furthermore, it must be that $s^* > \underline{s}$, for suppose not. Then market clearing (4) implies that $k_0(\sigma) > 0$ for all σ (task output must be strictly positive due to the INADA properties of the Cobb-Douglas production function). By (7), for some (s,σ)

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

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 (2).

(b) If $k_0(\sigma) > 0$, then by the zero-profit condition (7)

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

and there is some σ' such that $n_1(s^*, \sigma') > 0$ and hence by (7)

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

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 (σ, σ') and $s \leq s^*$ such that $k_0(\sigma) > 0$ and $n_0(s, \sigma') > 0$ we have by (7),

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

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 σ_0^*.

Similarly, for any (σ, σ') and $s \geq s^*$ such that $k_1(\sigma) > 0$ and $n_1(s, \sigma') > 0$, we have by (7),

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

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 α and since $s > s_K$. This establishes existence of σ_1^*.

Now, it must be that $\sigma_0^* < \sigma_1^*$, for suppose not. If $\sigma_0^* > \sigma_1^*$, then there exist (s, σ) such that $k_0(\sigma) > 0$, $k_1(\sigma) = 0$, $n_0(s, \sigma) = 0$, and $n_1(s, \sigma) > 0$. By (7),

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

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 (2). If $\sigma_0^* = \sigma_1^*$, then similar arguments lead to $\alpha(s, \sigma) = 1$, which also contradicts (2). ∎

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

A.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 $\widehat{M}'(\overline{s}) < M'(\overline{s})$, then $\widehat{M}(s) < M(s)$ for all $s \in [\max\{s^*, \widehat{s}^*\}, \overline{s})$. For suppose that $\widehat{M}'(\overline{s}) < M'(\overline{s})$ and that there exists some $s' \in [\max\{s^*, \widehat{s}^*\}, \overline{s})$ such that $\widehat{M}(s') \leq M(s')$. Then there exists some $s'' \in [s', \overline{s})$ such that $\widehat{M}(s'') = M(s'')$, $\widehat{M}'(s'') \geq M'(s'')$, and $\widehat{M}(s) > M(s)$ for all $s \in (s'', \overline{s})$. 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,

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

where

$$\omega(s; M) \equiv \exp\left[\int_s^{\overline{s}} \frac{\partial}{\partial z} \log \alpha(z, M(z)) dz\right]. \tag{27}$$

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

Plugging (27) into (10), we obtain

$$\frac{M'(\overline{s})}{M'(s)} = \omega(s; M)\frac{v(\overline{s})}{v(s)}. \tag{28}$$

Therefore,

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

By the above arguments, the right side of the last equation is larger than one, so that we must have $\widehat{M}'(\overline{s}) > M'(\overline{s})$, 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 $\widehat{M}'(\overline{s}) \geq M'(\overline{s})$.

By (22) and (28),

$$\frac{V(s^*)}{\overline{\sigma} - \sigma_0^*} \times \frac{M'(\overline{s})}{\omega(s^*; M)} = \frac{\beta v(\overline{s})}{1 - \beta}. \tag{29}$$

This together with (20), implies

$$\frac{A_K \alpha(s_K, \sigma_0^*) K}{\beta(\sigma_0^* - \underline{\sigma}) + (1 - \beta)(\sigma_1^* - \underline{\sigma})} \times \frac{M'(\overline{s})}{\omega(s^*; M)} = \frac{v(\overline{s})}{1 - \beta}. \tag{30}$$

Suppose that $\widehat{s}^* \geq s^*$. Then (29) implies that $\widehat{\sigma}_0^* < \sigma_0^*$, while (30) 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), $\widehat{s}^* > s^*$,[30] so it must be

[30] 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 α, 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 σ_1^* raises the ratio $\alpha(s_K, \sigma_1^*)/\alpha(s^*, \sigma_1^*)$ since $s_K < s^*$. The rise in σ_0^* raises 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 $\widehat{M}'(\overline{s}) < M'(\overline{s})$. 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 $\widehat{M}'(\overline{s}) > M'(\overline{s})$, which we have shown implies $\widehat{M}(s) < M(s)$ and, by (28), $\widehat{M}'(s) > M'(s)$ 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,

$$\frac{\widehat{w}(s)}{\widehat{Y}} = \frac{1-\beta}{\mu}\frac{\widehat{M}'(s)}{v(s)} > \frac{1-\beta}{\mu}\frac{M'(s)}{v(s)} = \frac{w(s)}{Y} \quad \forall s \in [\widehat{s}^*, \overline{s}].$$

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

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

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

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

By (10) and (19), this implies $\beta\widehat{\sigma}_0^* + (1-\beta)\widehat{\sigma}_1^* < \beta\sigma_0^* + (1-\beta)\sigma_1^*$.

Now observe that if $\widehat{M}(s) < M(s)$ then $\omega(\widehat{s}^*; \widehat{M}) < \omega(s^*; M)$ since also $\widehat{s}^* > s^*$. By (29), we must have $\widehat{\sigma}_0^* < \sigma_0^*$. But this means that (30) can only hold if also the total labor share has decreased, $\beta\widehat{\sigma}_0^* + (1-\beta)\widehat{\sigma}_1^* > \beta\sigma_0^* + (1-\beta)\sigma_1^*$, 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 knowledge-intensive. Formally, if $\beta\widehat{\sigma}_0^* + (1-\beta)\widehat{\sigma}_1^* < \beta\sigma_0^* + (1-\beta)\sigma_1^*$, then $\widehat{\sigma}_1^* < \sigma_1^*$. For suppose that $\beta\widehat{\sigma}_0^* + (1-\beta)\widehat{\sigma}_1^* < \beta\sigma_0^* + (1-\beta)\sigma_1^*$, 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 knowledge-intensive, 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 $s' \in (\max\{s^*, \widehat{s}^*\}, \overline{s}]$, 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

the LHS further. Therefore, s^* must increase.

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. ∎

A.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 α. 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 $\beta(\overline{\sigma} - \sigma_0^*) + (1-\beta)(\overline{\sigma} - \sigma_1^*)$. As $\widehat{\sigma}_1^* > \sigma_1^*$, the result is immediate if $\widehat{\sigma}_0^* \geq \sigma_0^*$. Then consider the case $\widehat{\sigma}_0^* < \sigma_0^*$. Rewrite (20) as

$$A_K \alpha(s_K, \sigma_0^*) K = \frac{\beta(\sigma_0^* - \underline{\sigma}) + (1-\beta)(\sigma_1^* - \underline{\sigma})}{\frac{\beta(\overline{\sigma} - \sigma_0^*)}{V(s^*)}}.$$

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. ∎

B A Model with Fixed Costs

We begin by simplifying the modeling of the task production process. Assume that the set of potential problems encountered in each task is given by $[0, \sigma]$. Moreover, suppose that machines and workers can only be employed in a given task if they can solve *all* problems in this interval. Thus, we abstract from training and design choices. Nevertheless, the concept of knowledge intensity is still present in the model and is captured by the parameter σ. The technologies for training workers and designing machines in the modified model are as follows. Intermediate firms must pay σ/s units of the final good to train a worker in training-intensive task σ, but face no learning cost in innate ability tasks. Maintaining the normalization that task-neutral productivity of workers equals one, we have that the marginal cost of employing labor is given by $w(s) + \sigma/s$.

To design a machine in a task with knowledge intensity σ, be it a training-intensive or an innate ability task, firms pay a one-off cost $\varphi\sigma$ and a variable cost $c_K\sigma$. Thus, the marginal cost of employing machines is $r/A_K + c_K\sigma/A_K$, where r is the rental rate of capital and A_K is task-neutral productivity of machines.

We assume that each task is produced by a single monopolistic firm.[31] In contrast, final good firms are perfectly competitive just as in the baseline version of the model. The final good production function is now

$$Y = \left[\int_{\underline{\sigma}}^{\overline{\sigma}} \left\{ \beta y_0(\sigma)^{\frac{\varepsilon-1}{\varepsilon}} + (1-\beta) y_1(\sigma)^{\frac{\varepsilon-1}{\varepsilon}} \right\} d\sigma \right]^{\frac{\varepsilon}{\varepsilon-1}},$$

with $\varepsilon > 1$. 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 the profits of the firm that supplies training-intensive task σ are given by

$$\pi_1(\sigma, s|N) = a_1(\varepsilon) \left[w(s) + \sigma/s \right]^{-(\varepsilon-1)}$$

if employing workers of type s, and

$$\pi_1(\sigma|K) = a_1(\varepsilon) \left[r/A_K + c_K\sigma/A_K \right]^{-(\varepsilon-1)} - \varphi\sigma$$

if employing machines, where $a_1(\varepsilon) \equiv \varepsilon^{-\varepsilon}(\varepsilon-1)^{\varepsilon-1}(1-\beta)^{\varepsilon}Y$. In innate ability tasks, the corresponding expressions are

$$\pi_0(\sigma, s|N) = a_0(\varepsilon) w(s)^{-(\varepsilon-1)}$$

and

$$\pi_0(\sigma|K) = a_0(\varepsilon) \left[r/A_K + c_K\sigma/A_K \right]^{-(\varepsilon-1)} - \varphi\sigma$$

with $a_0(\varepsilon) \equiv \varepsilon^{-\varepsilon}(\varepsilon-1)^{\varepsilon-1}\beta^{\varepsilon}Y$. Unlike in the baseline model, incentives for employing machines depend both on knowledge intensity and training intensity. This is because of a market size effect that is present whenever the share of innate ability tasks β is different from one half.

The equilibrium assignment of machines and labor to intermediate firms is qualitatively the same as in the baseline model if the marginal costs of employing machines are lower than those of employing workers. In particular, if for all s, $w(s) > r/A_k$ and $1/s > c_K/A_K$, and if $\underline{\sigma}$ is close to zero, then $\pi_0(\underline{\sigma}, s|N) < \pi_0(\underline{\sigma}|K)$ and $\pi_1(\underline{\sigma}, s|N) < \pi_1(\underline{\sigma}|K)$ for all s. Thus, the least knowledge-intensive innate ability and training-intensive tasks are performed by machines. Now

[31] 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.

observe that profits of firms employing labor approach zero, but stay strictly positive, as σ goes to infinity. In contrast, profits of firms employing machines will be negative for sufficiently large σ due to the fixed cost. Therefore, if $\overline{\sigma}$ is large, then there exist σ_0^* such that innate ability tasks with $\sigma \leq \sigma_0^*$ ($\sigma > \sigma_0^*$) are performed by machines (workers). Similarly, there exists such a marginal training-intensive task σ_1^*. If β is not too large, then $\sigma_1^* > \sigma_0^*$, so that machines are more widely adopted in training-intensive tasks. We have thus established the conditions under which technology adoption in the model with fixed costs follows the same patterns as in the baseline model (Lemma 2, part b).

Now consider the assignment of skill types to training-intensive tasks. If the firm supplying training-intensive task σ employs type s in equilibrium, then its profits are equal to $\pi(\sigma, s|N) = a_1(\varepsilon)\,[w(s) + \sigma/s]^{-(\varepsilon-1)}$. For this to be optimal, the first-order-condition

$$w'(s) - \sigma/s^2 = 0$$

and the second-order condition

$$w''(s) + 2\sigma/s^3 > 0$$

must hold. For firms supplying more-knowledge-intensive tasks to hire more highly skilled workers, it must be that $ds/d\sigma > 0$. It is easy to check that this condition is satisfied under the first- and second-order conditions above. Thus, the matching function is increasing and there is positive assortative matching as in the baseline model. Since the wage function is increasing, there must exist an s^* such that all workers with $s < s^*$ ($s \geq s^*$) are assigned to innate ability tasks (training-intensive tasks). The assignment of skill types to tasks is thus equivalent to that in the baseline model (Lemma 2, part a, and Lemma 3).

To solve for the matching function, follow similar steps as in the derivation of (10) to obtain the differential equation

$$M'(s) = \left(\frac{\varepsilon}{(\varepsilon-1)(1-\beta)}\right)^{\varepsilon} \times \frac{v(s)\,[w(s) + M(s)/s]^{\varepsilon}}{Y}.$$

Together with the FOC (setting $\sigma = M(s)$) and the boundary conditions $M(s^*) = \sigma_1^*$ and $M(\overline{s}) = \overline{\sigma}$, one can solve for the matching function, given a guess for s^*, σ_1^*, and Y. The model is closed by the usual market clearing conditions and the no-arbitrage equations $\pi(\sigma_0^*, s^*|N) = \pi(\sigma_0^*|K)$, $\pi(\sigma_1^*, s^*|N) = \pi(\sigma_1^*|K)$, and $w(s) = w(s^*)$ for all $s < s^*$.

Table C.1: Measuring Training Requirements Based on SVP and Job Zones

	SVP	*Job Zone*	*Training time*
1	short demonstration	1	1.5 months
2	up to 30 days		
3	30 days to 3 months		
4	3 to 6 months	2	7.5 months
5	6 months to 1 year		
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		

C 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 6.2) 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 C.1 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 C.1 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 C.2 lists the twenty least and most training-intensive occupations (using David Dorn's classification) in 1971. Table C.3 does the same for 2007. Table C.4 lists the twenty occupations experiencing the largest declines and increases in training requirements.

Table C.2: Least and Most Training-Intensive Occupations, 1971

Occupation (occ1990dd grouping)	Training requirements in years (1971)
a) least training-intensive	
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
b) most training-intensive	
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 C.3: Least and Most Training-Intensive Occupations, 2007

Occupation (occ1990dd grouping)	Training requirements in years (2007)
a) least training-intensive	
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
b) most training-intensive	
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 C.4: Largest Decreases and Increases in Training Requirements, 1971-2007

Occupation (occ1990dd grouping)	Change in training requirements (years) 1971-2007	Training requirements in 1971 (years)
a) largest decreases in training requirements		
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 refigeration 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
b) largest increases in training requirements		
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