How It's Made: A General Theory of the Labor Implications of Technology Change *

Christophe Combemale*, Laurence Ales, Erica R.H. Fuchs, Kate S. Whitefoot

Carnegie Mellon University

This Version: October 7, 2021

Abstract

This paper develops a general theory relating technology change and skill demand, capable of rationalizing the labor impacts of various technology changes since the 19th century. Performers (human or machine) face stochastic issues that must be solved in a given time to complete tasks. Firms choose how production tasks are divided into steps, the rate at which steps need to be completed, and the type of performer assigned to a step. Performers differ in the breadth of issues they can solve (generality) and in their tolerance for working at higher rates. Human performers tend to be generalists with low rate-tolerance. Machine performers tend to be specialists less sensitive to rate. Central to the theory are the cost of fragmenting tasks into smaller steps, the cost of allocating performers to multiple steps, and the negative relationship between step complexity and the rate of completing that step. We derive the cost-minimizing division of tasks and level of automation of production and the demand for workers of different skills that those conditions create. Our theory predicts that the division of tasks under increased complexity is skill polarizing; automation is skill polarizing at lower production volumes and upskilling at higher volumes; and that parts consolidation increases the demand for mid-level skills. We find counterparts to the theory across a range of industrial contexts and time periods, including the Hand-Machine Labor Study covering mechanization and process improvement at the end of the 19th century, automotive body assembly, and emerging technological changes in optoelectronic semiconductors used for communications.

JEL Codes: *D24, E23, N6*
Keywords: *Skill-Biased Technology Change, Division of Labor, Automation.*

*E-mail: ccombema@andrew.cmu.edu (Combemale). We thank Brian Kovak and Ersin Korpeoglu for helpful discussions as well as seminar participants at the U.S. Census, the 2021 Consortium on Competitiveness and Cooperation Doctoral Conference, the 2021 Industry Studies Association Annual Conference, and at the 2021 SED Meeting in Minneapolis. We also thank Zachary Leventhal for invaluable help with the digitization of the Hand and Machine Labor Study. This research is supported by grants from CMU's Manufacturing Futures Initiative, the National Bureau of Economic Research, the National Science Foundation, the American Institute of Manufacturing Photonics and the Russell Sage Foundation.

1 Introduction

Technology change plays an important role in labor markets, impacting inequality and the wage returns to skill. At the same time, technology has impacted workers in different manners over time. For example, the adoption of the factory system and machinery during the nineteenth century led to de-skilling (Goldin and Katz, 1998; Hounshell, 1985) while the automation of routine tasks during the 1970s-1990s led to up-skilling (Autor, Levy, and Murnane, 2003).[1] While the literature has provided compelling evidence and explanation of these patterns in terms of substitutability between the capital that embodies a technology and worker skill, it does not explain why these differences in substitutability exist. This paper fills that gap.

In this paper we develop a general theory with the goal of understanding *why* different technologies impact workers differently, by describing how the demand for workers' skill level is endogenously determined and deriving five dimensions on which technological change can affect this process. We formalize the theory in a structural model capable of analyzing operations-level production data, and we provide empirical counterparts to the theory showing how different technological changes can differentially affect skill demand. The model we develop is rich enough to incorporate insights from information theory, computer science and industrial engineering concerning the production of goods and services, and yet tractable enough to derive implications of technology change for the division of production tasks, automation and skill demand.

The starting point of the model is the set of tasks that must be completed to make a product or a service. To minimize the cost of producing at a given volume, a firm must choose how to divide this set of tasks into production steps. The firm also chooses the performer type for each step (human or machine), and the rate of production for each step. A basic feature of the model is that the difficulty of a step is increasing in the number of tasks (the length of step) and in the rate at which the step needs to be completed. The degree to which the difficulty of a step is impacted by the number of tasks or the rate of completion is specific to the type of performer: humans are less sensitive to the number of tasks than machines (more general than machines), but more sensitive to rate. In deciding the division of production, the firm faces a trade-off. More difficult steps require a more able, thus more costly performer. This mechanism provides incentive for smaller steps. On the other hand, division of two sequential tasks incurs fragmentation costs, providing an incentive for longer steps. The firm must also take into account excess performer capacity, either by allowing a performer to be idle or by reallocating the performer to a different step. Reallocation incurs a performer specific divisibility cost. This cost is higher for machines than humans. Within the context of our model, technological change can be described in terms of how it alters five dimensions: 1) the overall complexity of a process, 2) the cost of dividing tasks in a process, 3) the sensitivity of performers to the rate of production, 4) the sensitivity of performers to the number of tasks in a step, 5) the cost of dividing performers among multiple

[1]The literature studying the impact of technology on workers and specifically the way in which different technologies differ is vast. As a starting point refer to: Caselli (1999), Bresnahan, Brynjolfsson, and Hitt (2002), Acemoglu and Autor (2011) Autor and Dorn (2013), Dinlersoz and Wolf (2018), Eden and Gaggl (2019), Acemoglu and Restrepo (2020).

steps.

We use the theory to characterize the impact of key technological changes on production and workers, with three main results. First, we identify conditions under which it is optimal for firms to divide production into smaller steps. We show that heterogeneous costs of dividing different tasks are necessary for heterogeneity in performer ability demand within a firm (skill demand in workers). We find that for division to occur, performer costs must be convex in the length of steps. This convexity occurs with sufficiently convex wages or with a sufficiently high production volume. From a historical perspective the optimality of division of labor under high volume rationalizes the adoption of the factory system and later assembly line in production contexts that did not experience significant variation in wages.

Second, we provide conditions under which it is optimal to automate a step. We find that two dimensions determine the choices of automation: the volume of production and the step length. Within these two dimensions our theoretical results identify a so-called *cone of automation*. Specifically, we find that at sufficiently low production volumes no automation is optimal, because the higher divisibility costs of machines lead firms to leave them idle, thus raising costs. At middle production volumes it is optimal to automate middle-length steps. This causes machines to substitute for middle skill workers, generating skill polarization. Short steps are not automated because they have high rates of work and hence low machine utilization, leading to high idling costs. Long steps are also not automated at middle production volumes. As steps increase in length, the cost of a machine performer increases faster than a human performer, because machine performers are less general than human performers. At high volumes, machine utilization is high even at high rates of work, and so only the longest steps are not economical to automate (substituting for low and middle skill workers). The cone of automation is a useful result for understanding the root causes of historical variation in the effects of automation (Goldin katz) and the more recent polarization of occupational demand (goos and manning).

For our third main result, we explore how changes in the division of tasks can affect skill demand and hence wages. We show that declining costs of dividing tasks (occurring during the initial phases of the industrial revolution) reduces the lower bound of skill demand. We also consider technologies that reduce fragmentation costs but increase process complexity, (such as modularization). We find that such technologies increase inequality between the highest and lowest wages by polarizing the upper and lower bounds of skill demand. This result also shows that technologies such as parts integration, reducing process complexity by eliminating opportunities to divide tasks, can reduce inequality between the highest and lowest wages.

We take our model to the data and provide empirical counterparts to key results of the theory. The model presented in the paper is rich enough to provide a tight linkage with production operations data. Up until recently, this type of data has rarely been used in the economic analysis of technology change. We use three sources of detailed operations data. The first dataset is the Hand and Machine Labor Study (Wright, 1898), covering mechanization and process innovations at the time of the Second Industrial Revolution (1870s to 1910s). This dataset covers 15,700 process steps for 671 products spanning mining, agricultural, manufacturing and transportation service products. The other two data sets are contemporary and novel (collected by some of the authors in the present paper), capturing in great detail the optoelectronic semiconductor component production and assembly contexts (Combemale, Whitefoot, Ales, and Fuchs, 2020) and the automotive body assembly context (Fuchs, Field, Roth, and Kirchain, 2008). The optoelectronic semiconductor data involves hand-collected shop-floor-level production data on five different design and production alternatives for a single data communications product. The data comes from extensive line observations, technical interviews (including skill assessments for each production step using the O*NET survey instrument) and operations data capturing the entirety of production at firms representing 42 percent of the industry's production volume. The automotive body assembly data contains detailed data on process flow from multiple major U.S. vehicle manufacturers and key inputs such as machine type and price as well as quantifiable engineering measures of process complexity (e.g. number of joins per step).

The detailed production data we use supports multiple empirical connections to our theory. We start by connecting fundamental assumptions of the model to empirical evidence from our contemporary datasets. We find evidence of trade-offs between the number of tasks in steps and the rate of operations (in automotive and optoelectronic semiconductor data). We also find evidence that the level of ability demand is increasing in the number of tasks in steps (in optoelectronic semiconductor data).

We find that the theory can rationalize patterns of substitution of machines for human workers during the second industrial revolution: we recover an empirical analog to the cone of automation directly from the production data in the Hand and Machine Labor study. We also show polarization of ability demand under automation in the optoelectronics context, consistent with the automation implications of our theory at middle production volumes.

We also show that our theory can explain historical and contemporary changes in the distribution of worker ability demand under different technological regimes for the cost of division of tasks. We show in the Hand and Machine Labor context that an increase in the division of tasks leads to polarization of the highest and lowest wages, consistent with predictions of the theory for the effects of technology changes at the time such as interchangeable parts. The theory is also consistent with our observations in optoelectronics that parts consolidation (reducing the divisibility of tasks but also process complexity) leads to the convergence of the domain of ability demand, with less demand for the highest and lowest ability and more at the middle.

Literature The idea that the division of labor is an important feature driving the demand for labor and productivity goes back at least to Adam Smith, with the famous pin-factory example (Smith, 1776). A small literature has analyzed when division of tasks should occur and what are the limits to the division of tasks. Smith himself argues that the degree of specialization is limited by market size, as small market sizes do not generate enough demand to support specialized firms. This insight is also supported by Stigler (1951). Other work has characterized the productivity returns of and limits to the division of labor. For example, in Becker and Murphy (1992); Yang and Ng (1998), a task is split across workers and the upper bound on the division of labor is given by coordination costs across teams. In our model this mechanism is captured by the costs of splitting different tasks from each other and assigning them to different workers.

This paper is related to the literature modeling the task content of production (Autor, 2013; Acemoglu and Restrepo, 2018a,b). Similar to this literature, we consider a job as a bundle of tasks and model the optimal assignment of a task to either a human or a machine. The emphasis of this literature (differently than ours) is in considering the long run effects of displacement of workers by capital. We take a broader approach to technology (going beyond automation) and emphasize the circumstances under which workers of different ability levels are displaced by other differently able workers or machines.

This paper relates to the literature on polarization of occupational demand (Goos, Manning, and Salomons, 2009; Acemoglu and Autor, 2011; Goos, Rademakers, Salomons, and Vandeweyer, 2019). This literature has identified aggregate changes in the occupational structure of advanced economies in the last few decades. Polarization refers to the fact that middle-wage occupations exhibit lower (or negative) growth relative to low and high paying occupations. This has been put forward as evidence of ICT-capital adoption replacing mid-level skills. With respect to this literature, our theory provides a micro-founded mechanism for these occupational changes. We show the condition in which automation is more likely to occur for mid-level skills (also considering when automation occurs for low-level skills). In addition, the data presented in this paper provides additional evidence of the polarization phenomenon being present when looking at workers within a plant.

This paper also relates the literature on the labor consequences of different forms of automation, from traditional mechanization (Goldin and Katz, 1998) to robotics (Graetz and Michaels, 2018; Acemoglu and Restrepo, 2020) to machine learning (Brynjolfsson et al., 2018). We do so by explaining how these and other technological changes affect task divisibility and may generate differential labor outcomes. For example in our theory, robotics offers more general performers than traditional mechanization, leading to more automation of high skill steps, while machine learning offers both greater generality and greater divisibility, which leads to more automation of high and low skill steps.

Our modeling strategy is related to the approach in Garicano and Rossi-Hansberg (2006), in which firms organize themselves into hierarchies. In their model problems of varying complexity are divided and assigned to different workers. As in our paper this approach creates an endogenous relationship between earnings and talent. Similarly to Garicano and Rossi-Hansberg (2006), we model production as generating a series of issues that need to be resolved. Differently from their paper, our theory allows for tasks to be arbitrarily divided, for different types of performers (human and machines), and for the production rate to be endogenously determined.

Our work is also related to the task-assignment literature.[2] The literature studies the optimal

[2] See for example, Rosen (1978); Costinot and Vogel (2010); Ales et al. (2015); Lindenlaub (2017); Ocampo (2018);

assignment of heterogeneous workers to jobs of varying complexity or composition. The bulk of the task-assignment literature is fairly general in the set of jobs and skills analyzed. This is expected as the scope of the analysis encompasses the entirety of the labor market. Most of the work in this area studies properties of the indirect production function over occupations inherited from the assignment problem. To make progress in this direction, strong assumptions on the primitives of the production function are needed. Our approach is closer to the original motivation of Rosen (1978)): we characterize the endogenous bundling and assignment of work activities.[3] In our model not only the assignment is endogenous but so also is the complexity of the job (determined by the set of tasks and the rate of production).

This paper aslo builds on the literature relating technology change to process and firm structure. The idea of fragmentation costs in this paper connects to past work on modularity and integration in product and process design (Baldwin and Clark, 2003; Baldwin, 2008): we extend these costs to motivate heterogeneity in production steps and introduce performer characteristics. Prior work connects organizational changes with technological change and skill demand (Caroli and Van Reenen, 2001; Bresnahan et al., 2002): our model allows skill demand effects of new technology to originate from substitution of performers within existing steps (a non-organizational change) but also from the reorganization of tasks (organizational change).

Layout The paper proceeds as follows. Section 2 motivates and describes the key ingredients in the model. Section 3 formalizes the model. In Section 4 we analyze the implications of the model. We establish conditions for the optimality of division of tasks; give the relationship between step complexity and optimal rate and ability demand; describe different patterns of automation and their conditions and describe implication of changes in fragmentation costs. Section 5 provides empirical counterparts on the main findings of this paper. Section 6 concludes.

2 Empirical Motivation

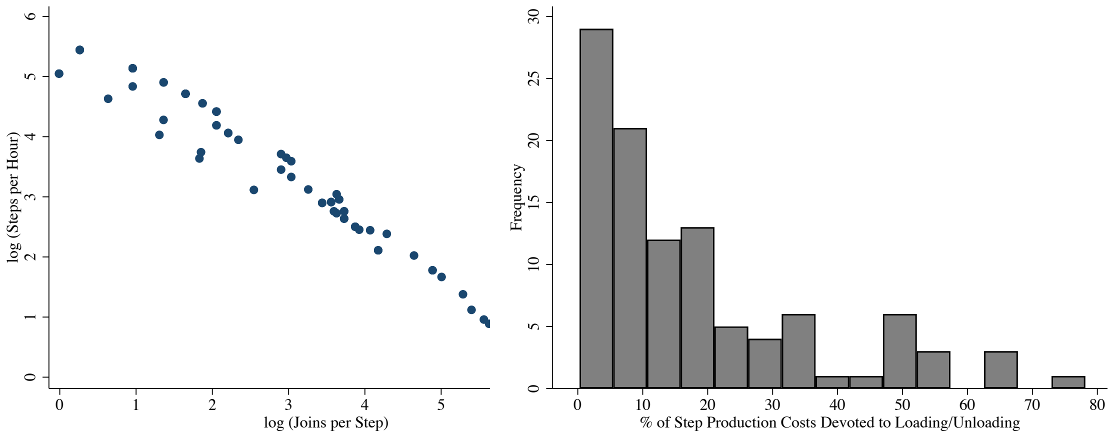
The starting point of the theory is the set of tasks that the firm must complete to produce a good, and the ability of a firm to divide production in multiple steps and assign these steps to either a human or machine performer. This feature is key as it will generate an endogenous demand for performers with a different ability level. Before formalizing the model, this subsection describes key features of technological change that have been analyzed in the economic and industrial engineering literature. These features will determine the key ingredients of the model.

The historical literature provides extensive examples of the importance of the division of tasks for early US manufacturing. For example, Hounshell (1985) and Womak, Jones, and Roos (1990) provide specific measurement for Ford automotive assembly plants.[4] They report that with

Haanwinckel (2020).

[3] We also study the effects of performer indivisibilities on differential returns to scale, a feature whose importance Rosen emphasized but did not include in the model.

[4] For modern examples of the benefits of division of tasks outside of manufacturing refer to Staats and Gino (2012).



(a) Speed Complexity Tradeoffs

(b) Fragmentation Costs

Figure 1: Motivation for model ingredients. For information on the data used for (a) refer to Fuchs et al. (2008). For information on the data used for (b) refer to Combemale et al. (2020).

the introduction of the moving assembly line around 1913, the average cycle time of a worker decreased from 2.3 to 1.2 minutes (the cycle time was 514 minutes before a fine division of tasks was introduced). At the same time the total amount of worker time per vehicle declined by 88 percent. Hounshell (1985) and Womak et al. (1990) also report that the demand for the skill of workers also changed during the move from craft production to factories to the adoption of the assembly line. In the time of craft production a worker was trained via lengthy apprenticeships on many aspects of automobile fabrication and assembly; however, by the time the assembly line was in full usage, the average training time for a worker was measured in minutes. This is an important ingredient of our theory: the fewer the tasks to perform, the easier the job for a worker.

The difficulty of completing a job is also driven by the overall time a worker or a machine has available to complete a task. The trade-off between measures of complexity and speed of execution has been extensively documented for both humans and machines.[5] The common denominator of these empirical regularities resides in the fact that any task requires information to be completed, and any operator has a limited bandwidth for such information (Shannon, 1948). Our own measurements confirm these regularities. In Figure 1a we display machine-level data from the automotive industry taken from Fuchs et al. (2008). In this case, it can be clearly seen how more complex part production (involving multiple joins per each step) is associated with an overall decrease in the number of completed steps per unit of time.[6]

Dividing production into ever smaller steps can introduce several benefits as described above. However, the division of production is not costless. When production is divided, one task in a

[5]See the work of Fitts (1954), Welford (1981) and MacKay (1982) for the case of human motor movements; For application to robotic systems refer to Lin and Lee (2013).

[6]The time-per-join varies across steps (steps with more joins tend to require less time per-join), so that the relationship between complexity and the rate of steps completed is not merely a linear function of the number of joins.

sequence is handled by a different performer from the next task. Transferring a work-in-progress from one performer to another takes time for both parties and creates errors. This phenomenon has been extensively studied, see for example Becker and Murphy (1992) and Baldwin (2008). Our own measurements illustrate the importance of these costs. In Figure 1b, we look at machine-level data from the optoelectronic semiconductor manufacturing industry taken from Combemale et al. (2020). A lower bound on the step fragmentation costs is the time devoted by the operator to loading and unloading a machine. For a large number of steps, this time-cost alone amounts to more than 10 percent of all step-wise production costs. Introducing fragmentation cost is also essential to understand the impact of a large number of technological developments. For example, a key development behind the growth of mass production is the introduction of exchangeable parts, which lowered the cost of splitting production across multiple workers and greatly increased productivity (Hounshell, 1985). Technological progress does not always lead to decreases in costs of splitting production. For example, parts integration in electronics reduces divisibility due to monolithic part integration (Combemale et al., 2020).

The previous costs are embodied in the technology used in production. An additional source of costs in dividing production tasks, and a final ingredient of the model, is incorporated in the cost of splitting performers across steps. Very short steps do not demand the full capacity of a performer, which introduces the possibility of a worker or a machine being under-utilized in production. Reallocating underutilized performers to other tasks is not costless, for instance incurring time to reconfigure machines or for workers to change tooling or position. Differently from the previous costs, these opportunity costs now depend on the total level of production (see Hopp and Spearman (2011) and Laureijs, Fuchs, and Whitefoot (2019) for an extensive analysis).

These ingredients give intuitive dimensions to the problem of the firm in dividing production tasks: the firm must trade off between the cost of complex steps and the cost of dividing tasks and performers. We formalize these dimensions in the following section.

3 Model

The description of the model proceeds in several steps. First we describe the nature of production in terms of tasks and steps. Then we introduce the difficulty associated with each step. The description of how human or machine performers differ follows and the problem of the firm concludes this section.

3.1 Tasks and Steps

A good or service is produced by executing a set of tasks. The set of tasks that need to be completed is described by the interval $\mathcal{V} = [0, \overline{v}]$ with $\overline{v}$ finite. Tasks are indexed by $\nu \in \mathcal{V}$. A task can be performed by a human or a machine. We codify this information with the indicator function: $o(\cdot) : \mathcal{V} \to \{m, h\}$. When $o(v) = h$ a human (or when $o(v) = m$ a machine) is performing task v. A consecutive group of tasks $\mathcal{S}_t \subseteq \mathcal{V}$ performed by either a single human or a machine is referred to as a step.[7] To define a step we introduce a series of $T \geq 1$ thresholds $\{s_t\}_{t=1}^{T}$ that split the set of tasks into steps. For all t we have $s_t \in \mathcal{V}$ and $s_T = \overline{v}$. T thresholds define T steps as follows: where $\mathcal{S}_t = (s_{t-1}, s_t]$ for $t = 2, \ldots, T$ and $\mathcal{S}_1 = [0, s_1]$. The type of performer in step t is defined with the indicator $o_t \in \{m, h\}$. For every step we associate a length $l_t = s_t - s_{t-1}$ for all $t = 2, \ldots, T$ and $l_1 = s_1$.

Remark 1. *Each task completed contributes to the final value of the good or service. Let Y be that value. Then denote with $y(\cdot) : \mathcal{V} \to \mathbb{R}_+$ the individual contribution of tasks to the value of the good or service. We then have:*

$$Y = \int_0^{\overline{v}} y(v) dv.$$

The indexing of the tasks and their relationship with value added is quite flexible. In general, our interpretation is that steps that include a larger measure of tasks than other steps, are also more complex. If $y(v)$ is a constant for all v, these more complex steps also contribute more to the value added of the good.

Associated with every consecutive pair of tasks there exists a fragmentation cost. This cost is paid by the firm whenever production is split into multiple steps, which are conducted by different performers. Fragmentation costs are characterized by the exact point at which a step ends, and by the type of performer executing the step. The costs are described by the function $f(\cdot, \cdot)$: $\mathcal{V} \times \{0, 1\} \to \mathbb{R}_+$. For a given production process split over T steps and executed by performers according to o_i, total fragmentation costs are then by: $\sum_{i=1}^{T} f(s_i, o_i)$. [8]

3.2 Jobs and Difficulty

Firms define each job, assigned to either a human or machine, by assigning the performer to a step of a particular length and by determining the rate at which the step needs to be completed.

[7] It is possible for multiple performers of the same type to be involved in the completion of a step, through parallelization or coordination. For instance in automobile assembly, a firm might employ multiple welders in parallel to meet a given production volume, with each welder independently performing the same tasks. Alternatively, the firm might require multiple performers to work simultaneously, such as when two workers lift a car door to place it into a vehicle frame Fuchs, Field, Roth, and Kirchain (2011). These cases are treated equivalently in the model, provided that the performer type and ability is the same. It is also common for performers of different types to work simultaneously on a specific unit (e.g. a human and a collaborative robot). In this case each performer is generally performing a different task: a human might be responsible for visual and cognitive tasks, while a robot may be responsible for strength-based tasks Vicentini (2021). In our model this case is described as separate steps (with fragmentation costs potentially incurred from the robot-human interaction). Because steps are defined by a performer, they do not distinguish between a human and a tool (and tooling may be part of the price of a performer); a tool makes a task easier for a human (or machine), while a machine performs a task (Frohm et al., 2008).

[8] As an extension to the model we develop above, fragmentation costs can be thought of as deriving from the costs of transitional steps that allow a performer to hand-off the output of their task to another performer. Formally, the cost of these additional steps is:

$$f(s_i, o_i) = E\left[\left(\sum_{j=1}^{N} (X_j)^{\rho_{o_i}}\right)^{\frac{1}{\rho_{o_i}}}\right]; \quad N = \lfloor f(s_i)/\lambda_f \rfloor; \quad o_i = h, m.$$

This formulation introduces the arrival of fragmentation issues λ_f as a primitive parameter. It also generates a higher fragmentation cost for machine performers than human performers whenever $\rho_m < \rho_h$.

These two dimensions define the different margins on which humans and machines have an advantage. As shown below, human performers (in general) have an advantage in the difficulty associated with step-length complexity, and machine performers (in general) have an advantage in the difficulty associated with the rate of completion of a step. The overall difficulty of a job for a human or machine performer is determined by these two dimensions as we explain next.

Complexity During the execution of a step, a performer needs to solve a number of issues that arise in production to complete the step. The complexity of each issue is modeled according to an i.i.d. random variable $X \in \mathcal{X} \subseteq \mathbb{R}_+$. We assume that all moments of X exist and are bounded. A key difference between a human and machine performer is the ability to solve closely related issues. For a typical machine the ability to solve any issue is independent of the ability to solve other issues. For a human performer the ability to solve an issue implies the ability to solve all easier issues. We formalize this distinction as follows. Given n issues X_i with $i = 1, \ldots, n$ the aggregate step-wide complexity is given by $\mathbf{X}(0|\rho) = 0$ and:

$$\mathbf{X}(n|\rho) = E\left[\left(\sum_{j=1}^{n}(X_j)^{\rho}\right)^{\frac{1}{\rho}}\right], \qquad n \geq 1.$$

The above equation is reminiscent of a CES production function with degree of substitutability ρ (and elasticity of substitution equal to $1/(1-\rho)$). In our formulation $\rho \in [1, \infty)$ represents a key property of a performer, we will refer to ρ as the *degree of generality*. Below we assume a human performer has a higher ρ than a machine.

Assumption 1. *Let ρ_h (ρ_m) be the degree of generality of a human (machine) performer. Then $\rho_h > \rho_m$.*

To understand the role of ρ, it is convenient to consider two extreme cases:

1. **Perfect Generalist** A perfect generalist is a performer with $\rho = \infty$. In this case: $\mathbf{X}^g(n) \equiv \lim_{\rho\to\infty} \mathbf{X}(n|\rho) = \max_{i=1,\ldots,n} X_i$, Let $X_{k:n}$ the k-th order statistic out of a sample of n draws of X. In this case the step-wide complexity for the perfect generalist is captured by $\mathbf{X}^g(n) = X_{n:n}$. This scenario captures the case in which only the most complex issues drive step-complexity for the performer, because solving an issue of given complexity implies the performer can solve all issues of lesser complexity.

2. **Perfect Specialist** At the opposite end, a perfect specialist is a performer with $\rho = 1$. This scenario captures the case in which each issue affects the step-wide complexity separately regardless of complexity. In this case $\mathbf{X}^s(n) \equiv \sum_{i=1}^{n} X_i$, so that the complexity of all issues contribute to the overall step-wide complexity.

To formally show the relationship between ρ and complexity, it is helpful to relate the definition of $\mathbf{X}(n|\rho)$ to an L^p norm. The result below follows from using Hermite-Hadamard inequalities for convex functions.

Lemma 1. *For all $n > 1$, if $\rho_h > \rho_m$ then $\mathbf{X}(n|\rho_h) < \mathbf{X}(n|\rho_m)$.*

Proof. In Appendix A. $\square$

We next look at the role of n in the definition of complexity. For all n and for $\rho \geq 1$ we have:

$$n^{1/\rho}(E[X^\rho])^{1/\rho} = \left(E\left[\sum_{j=1}^{n}(X_j)^\rho\right]\right)^{\frac{1}{\rho}} \geq E\left[\left(\sum_{j=1}^{n}(X_j)^\rho\right)^{\frac{1}{\rho}}\right] = \mathbf{X}(n|\rho),$$

with equality holding when $\rho = 1$ or $n \leq 1$. As the number of issues increases, the complexity of a step for human and machine performers increases at a different rate. To see this, consider the case for large n. Let $\mathbf{S}_n(\{X_i\}_{i=1}^n) = \sum_{i=1}^n (X_i)^\rho$ and $\overline{\mathbf{S}}_n = E[S_n] = nE[X^\rho]$. We then have from Proposition 2 in Biau and Mason (2015) that:

$$\lim_{n\to\infty} \mathbf{X}(n|\rho) \approx E\left[\overline{\mathbf{S}}_n^{1/\rho} + \frac{1}{2}\frac{1-\rho}{\rho^2}\overline{\mathbf{S}}_n^{1/\rho-2}\left(\mathbf{S}_n - \overline{\mathbf{S}}_n\right)^2 + \cdots\right] \approx n^{1/\rho}(E[X^\rho])^{1/\rho}. \tag{1}$$

From (1) we see that $\mathbf{X}(n|\rho)$ increases more quickly with n for lower values of ρ. Finally, $\mathbf{X}(n+1|\rho) - \mathbf{X}(n|\rho)$ is decreasing in n. This last observation is the basis for a concave relationship between step length and step complexity defined below.

Remark 2. *The difficulty of work originates from an interaction between tasks and the type of performer (Campbell, 1988). Research highlights how, in general, humans are better able to solve a wide variety of issues than machines (see for example Wickens, Hollands, Banbury, and Parasuraman 2015) and experience a smaller increase in errors as complexity increases. When it is possible to divide complex work into many less complex parts, this human advantage is reduced, and machines can compete with humans in terms of low error rates. Humans experience sharp increases in their rate of failure as they are made to perform the same work faster; while machines are not immune to this effect, they typically outperform humans in terms of the error-effect of repeating simple tasks faster. As an example of this trade-off, an industrial robot that can reliably perform its tasks can do so at much higher rates than a human, but would need to be reprogrammed and refitted to perform a different set of tasks, while a human could complete either set (Korsah et al., 2013).*

Issue Arrival Just as the magnitude of issues is uncertain, so is the number of issues that need to be solved in order to complete the step. To capture this feature we model issues as a compound Poisson process.[9] Issues arise according to a Poisson process with intensity λ. So that the probability of n issues arising in a step of length l is given by:

$$P_n(l) = \frac{(\lambda l)^n}{n!}e^{-\lambda l}.$$

The parameter λ governs the relationship between step length (l) and the expected number of issues denoted by $N(l) = \lambda l$. The performer-specific expected complexity (or simply complexity

[9]The modeling of difficulty with $\rho = 1$ becomes a version of the Cramer-Lundberg model. See also, Cai (2014).

henceforth) of solving the step is given by:

$$c(l|\rho) = \sum_{n=0}^{\infty} P_n(l)\mathbf{X}(n|\rho). \tag{2}$$

Complexity of a step inherits properties of $\mathbf{X}(n|\rho)$. The following lemma summarizes key properties of complexity used later in paper.

Lemma 2. *The function $c(l|\rho)$ is: (i) strictly increasing and (ii) strictly concave in step length l.*

Proof. In Appendix A. □

To fix intuition, it is helpful to go back to the case of performers being either perfect generalist or perfect specialist and see how different performer characteristics impact step complexity.

Example 1 (A Solved Case)**.** *This example derives a closed form equation for the complexity level of a step for the case of a perfect generalist and perfect specialist. Assume that each X_i is uniformly distributed in $[0,1]$. We then have that the expected value for $X_{n:n}$ is given by:*

$$E[X_{n:n}] = \frac{n}{n+1}.$$

Since the number of issues and their complexity are assumed independent of each other, we have that the expected total difficulty to complete steps of length l_i by a perfect generalist is given by:

$$c(l_i|\infty) = \sum_{n=0}^{\infty} \frac{n}{n+1}\frac{(\lambda l_i)^n}{n!}e^{-\lambda l_i} = \frac{1}{e^{\lambda l_i}\lambda l_i} + \frac{\lambda l_i - 1}{\lambda l_i}.$$

From the above (and as proved previously), it is easy to see directly that $D(l_i|\infty)$ is increasing and strictly concave in l_i for all λ. For a pure specialist, we have:

$$c(l_i|1) = \sum_{n=0}^{\infty} \frac{n}{2}\frac{(\lambda l_i)^n}{n!}e^{-\lambda l_i} = \frac{\lambda l_i}{2}.$$

In contrast to $c(l_i|\infty)$, $c(l_i|1)$ is linear in step length.

Rate & Difficulty We now consider the second key characteristic of a job: the rate at which it is performed. A firm may choose the rate at which a performer must complete the tasks in a step. A higher production rate increases performer output per unit time but also raises the overall difficulty of a step. To proceed, we need to determine the unit of time. For simplicity, we will normalize time so that a unit of time corresponds to a work shift, which is exogenous to the model. Next consider the rate in terms of the number of repetitions of a step per unit time denoted by $r \geq \underline{r} = 1$.

Having determined the complexity of a step (c) and the rate at which the performer executes the step (r), we can now determine the overall difficulty of a step with these characteristics.

Step difficulty is generated by an aggregator function $D : \mathbb{R}^2 \to \mathbb{R}$. A step with complexity $c(l|\rho)$ performed by a performer of type $o = h, m$ with rate r is associated with a difficulty $D\left(c(l|\rho), r|o\right)$. We assume the following for the difficulty function D:

Assumption 2. *The function D is: increasing in both arguments. Linear and unbounded with respect to the first argument (complexity), strictly convex with respect to the second argument (rate). The function D is differentiable in both arguments. Denote with D'_r the derivative of D with respect to r, we assume:*

$$D'_r(c, r, |h) > D'_r(c, r, |m); \qquad D''_r(c, r, |h) > D''_r(c, r, |m), \qquad \forall\, c > 0, r \geq \underline{r}; \tag{3}$$

$$D(c(l|\rho_h), \underline{r}|h) < D(c(l|\rho_m), \underline{r}|m), \qquad \forall\, l > 0. \tag{4}$$

Equations (3) and (4) formalize the differences between a human and machine performer with respect to sensitivity to rate. A step assigned to a human performer requires lower difficulty at low rate with respect to a machine performer. As the rate of the step grows, eventually the difficulty for a human performer overtakes the one of a machine performer. The functional form D allows a trade-off between length and rate for a constant difficulty level. Totally differentiating $D\left(c(l|\rho), r|h\right)$ and keeping a constant difficulty level we get:

$$\frac{dr}{dl} = - \cdot c'_l(l|\rho) \cdot D'_l / D'_r. \tag{5}$$

An important property of the D function is the sensitivity with respect to rate r, which we define as:

$$\sigma = 1 + r\frac{D''_r}{D'_r}. \tag{6}$$

The value of σ controls the sensitivity of difficulty to the rate. An example of a functional form that satisfies Assumption 2 is:

$$D\left(c(l|\rho), r|o\right) = c(l|\rho) \cdot (\underline{c} + r^{\varsigma}), \tag{7}$$

with $\underline{c} > 0$ and $\varsigma > 1$ to satisfy Assumption 2. The above specification features a lower bound on the difficulty which is independent of r. In the functional form given by (7) we have that $\sigma = \varsigma$.

3.3 Performers

So far we have discussed two key differences between performers: ρ and σ. The former determines the ease with which a performer addresses problems of increased complexity. The latter summarizes the tolerance of a performer to an increase in rate. In general, these characteristics vary between performer types (human vs machine) and among performers of the same type. For example, different machines can be characterized by their level of generality and rate-sensitivity. Humans may also differ from each other along these dimensions.

The use of the aggregator D defined in the previous Section implicitly assumes that performers are heterogeneous along a single-dimensional ability level (denoted with a). When assigning an operator to a step, the ability level of the performer needs to be commensurate with the difficulty of the step. Formally, for a performer of type $o = h, m$ with degree of generality ρ is capable of executing a step of length l with rate r if $a > D\left(c(l|\rho), r|o\right)$.[10]

Divisibility The final dimension characterizing performers is performer divisibility. Performers vary in the degree to which they can divide their time and reallocate their effort. A highly divisible human performer is able to complete additional tasks once the initial tasks associated with their job are completed. For example, a human computer programmer can quickly switch to answering emails once their programming tasks are completed. This performer therefore is not idle even when they can finish their tasks quickly (r is high), but can be reallocated to other productive tasks. In contrast, a robotic welding machine cannot switch to other tasks when the welding tasks are completed. The firm must pay for the performer (the rental price of capital in this case) even when they are idle.[11] Unlike ρ and σ, the degree of divisibility is influenced not only by the type of performer but also by exogenous policy such as minimum shift labor laws.[12]

In general there are two additional types of indivisibility of performers: the minimum time a performer can be allocated to the task, and the incremental amount that a performer can allocate to a task above the minimum (for example, a worker might work a minimum of four hours with hourly increments). We focus on the minimum time a performer can be allocated to the task. When encoding this restriction we assume that any higher rate of work provides no benefit. This restriction is summarized in the function: $g(R, r) : \mathbb{R}_+ \times \mathbb{R}_{\geq 1} \to \mathbb{R}_+$.[13]

The function g takes into account the rate at which a step is performed r and adjusts the performer step-cost accordingly. A higher r denotes a shorter performer time devoted to the step thus a lower performer cost for the step. The function g also takes into account the number of products produced R. The reason for this dependency is due to the fact that the benefit of raising r depends on the number of products to be processed and on the ability to reallocate the performer to a different task. To fix ideas we give two examples taken from Hopp and Spearman 2011:

1. **Perfectly divisible performer.** In this case we have:

$$g^{div}(R, r) = \frac{1}{r}.$$

In this case any increase in r translates into a proportionate reduction in the costs associated

[10] In our model the ability of a worker is single dimensional. We abstract from explicitly modeling workers characterized by multidimensional abilities. Refer to Lindenlaub (2017) and Ocampo (2018) for work in this area.

[11] The inability to fully use the capacity of a performer is a common concern in the systems engineering literature. Refer to Hopp and Spearman (2011) for an extensive analysis.

[12] The divisibility of performers can also be affected by institutional and organizational constraints. For example Schmitz and Teixeira (2008) when analyzing the Brazilian iron ore industry, document the productivity impact of organizational changes within the firm; specifically, they document how allowing repair staff to perform repairs outside their job classification (increasing the divisibility of the performer) increased labor productivity of these workers.

[13] Restriction on the incremental divisibility of workers would also be encoded in g. For analytical tractability we do currently do so.

with the performer completing the assigned step. For this case, cost reductions from higher r are independent of R.

2. **Indivisible performer**. This is the case of a performer that cannot be reallocated to a different task when idle. For this type of performer we have:

$$g^{ndiv}(R,r) = \frac{1}{R}\left\lceil \frac{R}{r} \right\rceil .$$

   In this case the gains from higher rate r are limited by the number of products produced (R).

In general the function $g(R,r)$ is assumed having the following properties:

Assumption 3. *For human ($i = h$) and machine ($i = m$) performers, the function $g^i(R,r)$ is such that:*

1. *For all R and i, there exists an $\bar{r}_i(R)$ such that $g^i(R,r) = g^i(R,\bar{r}_i(R))$ for all $r \geq \bar{r}_i(R)$. In addition $\lim_{R\to 0} \bar{r}_i(R) = 0$;*
2. *$\lim_{R\to\infty} g^i(R,r) = 1/r$;*
3. *If $r > r'$ then $g^i(R,r) \leq g^i(R,r')$ for all R;*
4. *If $R' > R$ then $g^i(R,r) > g^i(R',r)$ for all $r > \bar{r}_i(R)$;*
5. *For all R, $\bar{r}_h(R) \geq \bar{r}_m(R)$.*

Condition 1 in the above Assumption formalizes the idea that above a certain level of rate there are no further returns. This insight is commonly represented in the engineering literature by assuming that performers are "dedicated" to a process or to a step, meaning that their unused capacity cannot be productively used elsewhere. We refer to $\bar{r}_i(R)$ as the specific minimum divisibility threshold for the performer. When $r > \bar{r}$, all output is produced with a single performer within the minimum time increment, and so increasing r further cannot reduce the costs associated with the performer. *Condition* 2 States that as the number of items grows, the minimum threshold becomes non-binding. Finally, *Condition* 5 encodes the idea that moving a human performer to a different task is easier than re-tasking a machine performer.

We can now determine the total cost of assigning a performer to a step. Total step-costs are determined by the ability-price of the performers, $w(a)$ for humans and $k(a)$ for machines, as well as the cost saving associated with increasing the rate in which a step is executed, $g(R,r)$. We have that the price of a performer to complete a step with ability a, rate r and total number of products produced R is given by:

$$p(a,r,R|o_i) = \begin{cases} w(a)g^h(R,r), & \text{if } o_i = h \\ k(a)g^m(R,r), & \text{if } o_i = m \end{cases} . \tag{8}$$

We assume the following for functions $w(\cdot)$ and $k(\cdot)$:

Assumption 4. *The functions $w(\cdot)$ and $k(\cdot)$ are: positive, strictly increasing and weakly convex.*

We now have all the model ingredients needed to define the problem of the firm.

3.4 Firm Optimization

The firm chooses how to subdivide the production process by choosing the number and positions of steps, and which performer to assign a given step. For each step, the firm also determines the required completion rate. We begin by taking the number of steps T as given and finding the cost minimizing step thresholds s_i, operator, o_i, ability, a_i and rate, r_i, for each step i.

$$C(R,T) = \min_{\{s_i\}_{i=1}^{T}, \{r_i, a_i, o_i\}_{i=1}^{T}} \sum_{i=1}^{T} p(a_i, r_i, R|o_i) + \sum_{i=1}^{T} f(s_i, o_i) \tag{9}$$

subject to:

$$l_1 = s_1; \quad l_i = s_i - s_{i-1}, \quad \forall\, i = 2, \ldots, T; \tag{10}$$

$$a_i \geq D(c(l_i|\rho_{o_i}), r|o_i), \qquad \forall\, i = 1, \ldots, T; \tag{11}$$

$$s_i \in [0, \overline{v}]; \qquad s_i \leq s_{i+1}; \qquad o_i \in \{h, m\}; \qquad r_i \geq \underline{r}, \qquad \forall\, i = 1, \ldots, T; \tag{12}$$

$$s_0 = 0; \qquad s_T = \overline{v}. \tag{13}$$

The two terms in (9) represent the performer and fragmentation costs associated with a given choice of T (and performer characteristics). The per-unit cost of producing R units is then determined by choosing the cost minimizing number of steps T.

$$C(R) = \min_{T \in \mathbb{N}_+} C(R,T). \tag{14}$$

We assume that firms take as given $w(\cdot)$ and $k(\cdot)$.

4 Analysis

We analyze the environment of the model in four sections. First we analyze the conditions for the division of tasks to occur, finding that division can be driven by the structure of wages or by production volume. Second we analyze the relationship between step length and ability demand, rate and wages. Third we analyze the conditions for firms to automate steps. We then show how these effects vary with production volume, showing polarization of skill demand at low volumes and up-skilling at high volumes. We conclude by analyzing the effect of changes in fragmentation costs on the division of production and the distribution of ability demand. We show first that variation in fragmentation costs over tasks is necessary for variation in demand for performer type and ability, then show how changes in process technology can affect the inequality between the highest and lowest wages.

4.1 The Structure of Production: Division of Tasks

The production problem described earlier provides a rich set of possibilities on how the structure of production can be organized. The organization is impacted by the cost associated with performers and with fragmenting production. In this section, we discuss conditions under which the firm finds it optimal to divide tasks across performers. For simplicity, we refer to human performers with wage rates w. However, the wage rate can simply be replaced with the rental price of capital to extend the discussion to the division of tasks among machines.

Since the price of performers is strictly increasing in their ability, it follows that constraint (11) binds at the optimum. This result implies that a necessary condition for production tasks to be divided into more than one step is the existence of at least one $0 < l < \overline{v}$ and r', r'' such that:

$$p\Big(D\left(c(\overline{v}|\rho_h), r^*|h\right), r^*, R\Big|h\Big) > p\Big(D\left(c(l|\rho_h), r'|h\right), r', R\Big|h\Big) + p\Big(D\left(c(\overline{v}-l|\rho_h), r''|h\right), r'', R\Big|h\Big). \tag{15}$$

Where r^* is the optimal r without any division of tasks. The above inequality is strict since fragmentation costs are nonzero. We explore two forces that lead firms to divide tasks. The first is the effect of convex wages. The intuition for why convexity of wages lead to fragmentation is straightforward. A sufficiently convex wage in ability makes it extremely expensive for a firm to hire a worker to execute a large non-fragmented step. Formally:

Proposition 1. *Suppose that $f(\cdot, h)$ is sufficiently low and that $w(\cdot)$ is sufficiently convex. Then division of tasks is optimal.*

Proof. If fragmentation costs $f(\cdot, h)$ are sufficiently low, then the condition described in (15) is also sufficient. Suppose that, by contradiction, for all $l \in (0, 1)$ and all r', r'' we have:

$$p\Big(D\left(c(\overline{v}|\rho_h), r^*|h\right), r^*, R\Big|h\Big) \leq p\Big(D\left(c(l|\rho_h), r'|h\right), r', R\Big|h\Big) + p\Big(D\left(c(\overline{v}-l|\rho_h), r''|h\right), r'', R\Big|h\Big).$$

To reach a contradiction, set $l = \overline{v}/2$, using (8) we have:

$$w(D\left(c(\overline{v}|\rho_h), r^*|h\right))g^h(R, r^*) \leq w(D\left(c(\overline{v}/2|\rho_h), r'|h\right))g^h(R, r') + w(D\left(c(\overline{v}/2|\rho_h), r''|h\right))g^h(R, r''),$$

setting $r' = r'' = r^*$ the above implies:

$$w(D\left(c(\overline{v}|\rho_h), r^*|h\right)) \leq w(D\left(c(\overline{v}/2|\rho_h), r^*|h\right)) + w(D\left(c(\overline{v}/2|\rho_h), r^*|h\right)). \tag{16}$$

If w is sufficiently convex, so will the function $\tilde{w}(l) = w(D\left(c(l|\rho_h), r^*|h\right))$, reaching a contradiction with equation (16). □

The previous result considered division of tasks as a way to reduce the cost of production for sufficiently convex wages, trading off against fragmentation costs. The notion of connecting division of tasks to increases in production efficiency dates to Adam Smith in the Wealth of Nations, in his discussion of the division of labor (Smith, 1776). Smith himself argues that the degree of specialization may also be limited by market size; we turn to this channel for division of labor next. The following Proposition sharpens the trade-off present between fragmentation costs (related to cross-step coordination) and the size of output (related to the size of the market). Formally:

Proposition 2. *Suppose that: $f(\cdot,h)$ is sufficiently low; R is sufficiently high. If $D'_r = \frac{\partial D}{\partial r}$ is sufficiently small (or $\overline{v}$ is sufficiently large), then division of tasks is optimal.*

Proof. If fragmentation costs $f(\cdot,h)$ are sufficiently low, then the condition described in (15) is also sufficient. Suppose that, by contradiction, for all $l \in (0,1)$ and all r', r'' we have:

$$p\Big(D\left(c(\overline{v}|\rho_h), r^*|h\right), r^*, R\Big|h\Big) \leq p\Big(D\left(c(l|\rho_h), r'|h\right), r, R\Big|h\Big) + p\Big(D\left(c(\overline{v}-l|\rho_h), r''|h\right), r'', R\Big|h\Big).$$

Set $l = \overline{v}/2$, using (8) we rewrite the above as:

$$w(D\left(c(\overline{v}|\rho_h), r^*|h\right))g^h(R,r^*) \leq w(D\left(c(\overline{v}/2|\rho_h), r'|h\right))g^h(R,r') + w(D\left(c(\overline{v}/2|\rho_h), r''|h\right))g^h(R,r''). \tag{17}$$

Set the values of $r' = r'' = \widehat{r}$ so that $D\left(c(\overline{v}|\rho_h), r^*|h\right) = D\left(c(\overline{v}/2|\rho_h), \widehat{r}|h\right)$. From (5) we get:

$$\widehat{r} = r^* + \frac{c(l|\rho)}{\partial l}\frac{D'_l}{D'_r}\frac{\overline{v}}{2}, \tag{18}$$

From Assumption 3, for R sufficiently high, we have $g(R,r) \approx 1/r$ for all r. We rewrite (17) evaluated at $r' = r'' = \widehat{r}$ as $\widehat{r} < 2r^*$. From (18) we see that this condition is violated for D'_r is sufficiently small or $\overline{v}$ is sufficiently large thus reaching a contradiction. □

4.2 Ability and wages

In this section we explore the demand for ability in the production process. The demand for ability is determined by the length and the rate at which a step is completed. Since the latter can be adjusted in the short-run, we first provide a result linking length and rate. Then we provide a condition linking the length of a step and the overall ability level. The relationship between length and ability is a key property of the model, because it underlies the ability demand effects of automation and of division of tasks. To simplify the analysis for some of our results we assume the following structure for $g(R,r)$:

Assumption 5. *For human ($o = h$) and machine ($o = m$) performers, $g^o(R,r)$ is given by:*

$$g^o(R,r) = \begin{cases} 1/r, & \text{if } r \leq \overline{r}_o(R) = R\overline{r}_o \\ 1/R\overline{r}_o, & \text{if } r > \overline{r}_o(R) = R\overline{r}_o \end{cases},$$

with $\overline{r}_h > \overline{r}_m$.

In the above we can see that $\overline{r}$ is the rate at which the total amount of time required to produce the quantity R is equal to the minimum time increment that can be allocated to a performer. We also assume that the ability-price for performers (either human or machine) is well-behaved so that a unique rate level emerges for any step length. The following assumption guarantees this:

Assumption 6. *The function* $\mathbf{w}$*, defined as:*

$$\mathbf{w}(x, o_i) = \begin{cases} \frac{xw'(x)}{w(x)}, & \text{if } o_i = h \\ \frac{xk'(x)}{k(x)}, & \text{if } o_i = m \end{cases},$$

is increasing for all $x > 0$.

The next proposition considers the impact of step length on the optimal rate and ability level for a constant performer type.

Proposition 3. *Suppose Assumptions 4, 5 and 6 hold. In addition, assume the following: (a) D is separable between its two arguments and (b)* $\sigma > rD'_r/D$ *for all* r*. Given two steps* i *and* j *with the same performer, denote with* r_i (a_i) *and* r_j (a_j) *the optimal choice for rate (ability) in step* i *and* j*. Then if* $l_i > l_j$*, we have that (i)* $r_i \leq r_j$*, (ii) if* $r_i, r_j \in (\underline{r}, R\bar{r}_h)$ *then* $r_i < r_j$*, and (iii)* $a_i > a_j$*.*

Proof. Suppose that $o_i = o_j = h$ (similar arguments follow for a machine performer). From (9), for step length l the choice for r solves:

$$\min_{\underline{r} \leq r \leq R\bar{r}_h} \frac{w\left(D\left(c(l|\rho_h), r|h\right)\right)}{r}. \tag{19}$$

We begin with the proof of *(i)*. Suppose by contradiction that $r_i > r_j$. The case in which $r_j = R\bar{r}_h$ is obvious as it is not optimal to increase rate in step i (doing so would raise the ability requirement of the step without benefiting from a lower usage of operator time). Consider now the case in which $r_j < R\bar{r}_h$. If $r_j > \underline{r}$ we then have the following first order condition (from Assumption 4, the second order condition is verified) holding for both steps:

$$\mathbf{w}\left(D(c(l_s|\rho_h), r_s|h)\right) = \frac{D(c(l_s|\rho_h), r_s|h)}{rD'(c(l_s|\rho_h), r_s|h)}, \qquad s = i, j. \tag{20}$$

Let function g be defined as follows:

$$g(r_s) = \frac{D(c(l_s|\rho_h), r_s|h)}{rD'(c(l_s|\rho_h), r_s|h)},$$

note that given Assumption (a) in the statement of the Lemma, function g is independent of step length. We then have that:

$$g'(r) = \frac{r(D'_r)^2 - DD'_r - rDD''_r}{(rD')^2},$$

given Assumption (b) in the statement of the Lemma, we have from (6), that $\sigma = 1 + r\frac{D''_r}{D'_r} > \frac{rD'_r}{D}$

so that $g'(r) < 0$ for all r and all l. Hence:

$$\frac{D(c(l_i|\rho_h), r_i|h)}{rD'(c(l_i|\rho_h), r_i|h)} < \frac{D(c(l_j|\rho_h), r_j|h)}{rD'(c(l_j|\rho_h), r_j|h)}.$$

With the above, we then reach a contradiction with condition (20) since given the contradicting assumption together with Assumption 6, we have that $\mathbf{w}\left(D(c(l_i|\rho_h), r_i|h)\right) > \mathbf{w}\left(D(c(l_j|\rho_h), r_j|h)\right)$.

The case with $r_j = \underline{r} < r_i$ can be analyzed in a similar fashion as the interior case above. In this case we have that the first order condition is given by:

$$\mathbf{w}\left(D(c(l_j|\rho_h), r_j|h)\right) \geq \frac{D(c(l_j|\rho_h), r_j|h)}{rD'(c(l_j|\rho_h), r_j|h)}.$$

The contradiction is reached in a similar manner as the previous case.

We next show that *(ii)*: if $r_i, r_j \in (\underline{r}, R\bar{r}_h)$ then $r_j > r_i$. In this case we have that (20) holds for both steps. Given the previous result, we need to rule out $r_i = r_j$. In this case we have $g(r_i) = g(r_j)$. However, by Assumption 6, $\mathbf{w}$ is strictly increasing, it follows that $D(c(l_i|\rho_h), r_i|h) = D(c(l_j|\rho_h), r_j|h)$ reaching a contradiction since $l_i > l_j$.

We next show *(iii)*: $a_i > a_j$. If $l_i > l_j$, from the previous result we have that $r_i \leq r_j$ with the inequality strict if $r_i, r_j \in (\underline{r}, R\bar{r}_h)$. Starting from this interior case we have (using the shorthand $\mathbf{w}_j = \mathbf{w}\left(D(c(l_j|\rho_h), r_j|h)\right)$.)

$$\mathbf{w}_j = g(r_j) < g(r_i) = \mathbf{w}_i, \tag{21}$$

where the two equalities are from (20) and the assumption of interior r_i, r_j. While the inequality follows from the properties of g discussed above together with $r_i < r_j$. From (21) we have that $D(c(l_i|\rho_h), r_i|h) > D(c(l_j|\rho_h), r_j|h)$. Since constraint (11) in the firm's optimization problem is binding, it then follows that $a_i > a_j$.

Next consider the cases that involve either $r_i = r_j = \underline{r}$ or $r_i = r_j = \bar{r}_h$. In this case $a_i > a_j$ follows from $l_i > l_j$. Finally, we need to consider separately two remaining cases. First let $r_j = \bar{r}_h$ and $r_i < r_j$. In this case we have:

$$\mathbf{w}_i \geq g(r_i) > g(r_j) \geq \mathbf{w}_j, \tag{22}$$

where the weak inequality originates from the first order condition of (19) and the possibility that the upper bound for r_j or lower bound for r_i might be binding. As before from (22) we conclude that $a_i > a_j$. The remaining case has $r_i = \underline{r}$ and $r_j \in (\underline{r}, R\bar{r}_h)$ so that $r_i < r_j$. This case proceeds as before noting that now the second weak inequality in (22) is now an equality. □

The conditions on D in the previous Proposition hold for the simpler example with D given by (7). In this case separability is immediate and we have that (recall $\sigma = \varsigma$ for this functional form): $\varsigma > \frac{\varsigma r^{\varsigma}}{\underline{c} + r^{\varsigma}}$ whenever $\underline{c} > 0$.

4.3 Automation

In this section, we describe the conditions under which a firm automates a step by choosing a machine performer rather than a human. We show how automation impacts labor demand by showing that step length (l) and production quantity (R) is a key determinant for the patterns of automation. This section proceeds in three parts. We first show the existence of an upper bound on the length of automated steps, then show the existence of a lower bound: these results establish a region of automation and give us the effect of automation on the distribution of human performer ability demand. We then show how the range of steps automated evolves with production quantity (R) specifically we show that if a step of a given length is automated for a given R it is also automated for all $R' > R$. This result is key in showing that the range of steps automated grows as does R. Combining the results of this section leads to a pattern of automation in (R, l) space as displayed in Figure 2.

In this Section we assume that fragmentation costs for machines are higher than for humans.[14]

Assumption 7. *For all $s_t \in \mathcal{V}$, $f(s_t, h) < f(s_t, m)$.*

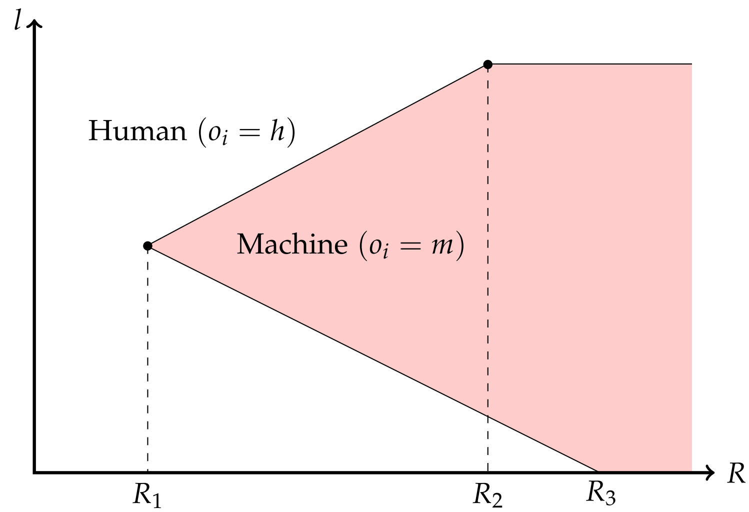


Figure 2: Automation patterns: volume and step length.

As we show below, if a step is long enough then it will not be automated, giving us the upper bound of a region of step lengths automated. This result is driven by the relatively higher generality (higher ρ) of humans.

Proposition 4 (Upper Bound on Automation). *There exists $\bar{l}$ such that $o_i = h$ for all i with $l_i > \bar{l}$.*

Proof. Suppose not, then for all $\bar{l}$ there exists a j with $\widetilde{l}_j > \bar{l}$ such that $o_j = m$. This implies that (recall $f(s_j, m) > f(s_j, h)$ from Assumption 7):

$$k(a_j^m)g^m(R, r_j^m) < w(a_j^h)g^h(R, r_j^h). \tag{23}$$

[14]Difficulties in machine-machine interactions transferring work in progress are well documented (Korsah et al., 2013), though this property is not essential for results of this section.

Since $k(\cdot)$ and $w(\cdot)$ are increasing, given Assumption 3 Part 1, the optimal r for either performer is always $\underline{r} \leq r \leq \bar{r}$. We then have. $D(c(l_j|\rho_h), r_j^h|h) = a_j^h \leq D(c(l_j|\rho_h), \bar{r}|h) \equiv \bar{a}(l_j)$, and $D(c(l_j|\rho_m), r_j^m|m) = a_j^m \geq D(c(l_j|\rho_m), \underline{r}|m) \equiv \underline{a}(l_j)$. Substituting the previous inequalities in (23) we have:

$$k(\underline{a}(l_j)) < w(\bar{a}(l_j))\frac{g^h(R,\bar{r})}{g^m(R,\underline{r})}. \tag{24}$$

For l sufficiently high, the probability of drawing a small number of issues is small. We can then approximate step difficulty in (2) as:

$$c(l_i|\rho) \approx \sum_{n=\underline{n}}^{\infty} P_n(l)\mathbf{X}(n|\rho)$$

with $\underline{n}$ sufficiently high. We can then use the approximation of $\mathbf{X}(n|\rho)$ in (1) and substitute into (24) so that we have:

$$k\left(D\left(\widetilde{n}_j^{1/\rho_m}E[X^{\rho_m}]^{1/\rho_m}, \underline{r}|\rho_m\right)\right) < w\left(D\left(\widetilde{n}_j^{1/\rho_h}E[X^{\rho_h}]^{1/\rho_h}, \bar{r}|\rho_h\right)\right)\frac{g^h(R,\bar{r})}{g^m(R,\underline{r})}.$$

Since $\rho_h > \rho_m$ (and $k = w$) the above is violated for n_j sufficiently large, reaching a contradiction. $\square$

The previous result is stated in terms of l sufficiently high for a given step. A similar result holds for any l if λ is instead sufficiently high. In both cases the step will feature a likely high number of issues.

We now consider the case of automation of small steps. This lower bound of automated step-lengths is driven by the lower divisibility of machines than humans, such that $\bar{r}_m < \bar{r}_h$ (defined in Assumption 5). The result holds as long as step difficulty for small steps is not affected by varying rate. As notation, in what follows let $r_h(l_i, R), r_m(l_i, R)$ be the optimal rate given the constraints $\bar{r}_h, \bar{r}_m$ and $r_h^*(l_i), r_m^*(l_i)$ the unconstrained optimal rate for human and machine respectively for a step of length l_i with output R.[15] Similarly o_i denotes the optimal choice of performer for step i. In the proposition that follows we assume that human wages are sufficiently low for low ability levels.

Proposition 5 (Lower Bound on Automation). *Suppose there exists a step i with l_i sufficiently small. Suppose also that $\lim_{c,r\to 0} w(D(c,r|h)) \leq k(D(c,r|m))$. Then if R is sufficiently low we have that $o_i = h$.*

Proof. Suppose not, then:

$$k(D(c(l_i|\rho_m), r_m|m))g^m(R, r_m) < w(D(c(l_i|\rho_h), r_h|h))g^h(R, r_h). \tag{25}$$

If R is sufficiently small we have that $r_m = \bar{r}_m(R) < \bar{r}_h(R) = r_h$. This implies that $g^h(R, r_h) <$

[15] That is $r_j(l_i, R) = \arg\min_{r \leq R\bar{r}_j}\left\{\frac{w(D(c(l_i|\rho_j),r|j))}{r}\right\}$; $r_j^*(l_i) = \arg\min_r\left\{\frac{w(D(c(l_i|\rho_j),r|j))}{r}\right\}$. When not a source of confusion, the dependency of r_j^* and r_j on l_i and R is omitted.

$g^m(R, r_m)$. From (25) it follows that $k(D(c(l_i|\rho_m), r_m|m)) < w(D(c(l_i|\rho_h), r_h|h))$. If l_i is sufficiently small, we then have that $c(l_i|\rho) \approx 0$; in addition, from Assumption 3 part 1 we have that $\lim_{R\to 0} r_j = 0$ for $j = h, m$. We then reach a contradiction with the assumption in the Proposition stating that $\lim_{c,r\to 0} w(D(c, r|h)) \le k(D(c, r|m))$.

□

The proof is straightforward and relies on the idea that for low R the advantage of a machine performer of operating at high rate is eliminated. This of course requires a minimum wage for workers that is sufficiently low. At the opposite end, with high R, we expect the presence of automation since in this case the optimal machine rate is higher than the human rate. For this to occur we need the symmetrical assumption on costs assumed in Proposition 5: $\lim_{c\to 0} [\lim_{r\to\infty} w(D(c, r|h))] \ge \lim_{c\to 0} [\lim_{r\to\infty} k(D(c, r|m))]$.

Between the upper and lower bounds of automated step lengths, automation is driven by the lesser sensitivity of machines to rate. For sufficiently short steps in which the constraint on machine rate is not binding, the lesser rate-sensitivity of machines allows them to achieve lower cost than humans. We next consider the optimality of automation as the product quantity increases. For the next proposition we keep the interval length fixed as we raise output R. This result is useful when comparing similar plants that operate at different scale. It also can provide insights on the optimal automation response of a plant faced with an increase in demand but not redesigning the entire production process. We have the following:

Proposition 6. *Suppose the g function satisfies Assumption 5. Suppose that $k(\cdot) = w(\cdot)$. Then if there exist i such that $o_i = m$ for a given R, then $o_i = m$ for all $R' > R$.*

Proof. Since step i is automated and $k(\cdot) = w(\cdot)$ it implies that:

$$\min_{r\le R\bar{r}_m} \left\{ \frac{w\left(D\left(c(l_i|\rho_m), r|m\right)\right)}{r} \right\} < \min_{r\le R\bar{r}_h} \left\{ \frac{w\left(D\left(c(l_i|\rho_h), r|h\right)\right)}{r} \right\}, \tag{26}$$

The proof proceeds by contradiction. Suppose that with $R' > R$ the step of length l_i is not automated. The contradicting assumption implies that:

$$\min_{r\le R'\bar{r}_m} \left\{ \frac{w\left(D\left(c(l_i|\rho_m), r|m\right)\right)}{r} \right\} \ge \min_{r\le R'\bar{r}_h} \left\{ \frac{w\left(D\left(c(l_i|\rho_h), r|h\right)\right)}{r} \right\}, \tag{27}$$

Let $\tilde{r}$ be defined as the rate such that $D(c(l_i|\rho_m), \tilde{r}|m) = D(c(l_i|\rho_h), \tilde{r}|h)$. This $\tilde{r}$ exists and is unique given conditions (3) and (4) in Assumption 2. For any $r < \tilde{r}$ we have that:

$$\frac{w\left(D\left(c(l_i|\rho_m), r|m\right)\right)}{r} > \frac{w\left(D\left(c(l_i|\rho_h), r|h\right)\right)}{r},$$

hence we have that for (26) to hold it must be the case that $\tilde{r} \le R\bar{r}_m$. Given (3) and (4), it also follows that $D'_r(c(l_i|\rho_m), r|m) < D'_r(c(l_i|\rho_h), r|h)$ and $D''_r(c(l_i|\rho_m), r|m) < D''_r(c(l_i|\rho_h), r|h)$ for all $r \ge \tilde{r}$. Since difficulty increases with respect to r at a faster rate for human relative to machine performers, we reach a contradiction with (27). □

The previous results look at patterns of automation when changing the length of a step or separately when changing the size of the output. We next consider the interaction between these two components as we will see below as the output size increases (R goes up) so does the region of step-lengths that is optimal to automate. The end result is a *cone of automation* as highlighted in Figure 2.

It is useful to define for a given R the maximum and minimum length of a step that will not be automated. The maximum length denoted with $\bar{l}(R)$ is the maximum l_i such that:

$$\min_{r \leq R\bar{r}_m} \left\{ \frac{w\left(D\left(c(l_i|\rho_m), r|m\right)\right)}{r} \right\} = \min_{r \leq R\bar{r}_h} \left\{ \frac{w\left(D\left(c(l_i|\rho_h), r|h\right)\right)}{r} \right\}.$$

The minimum length $\underline{l}(R)$ is similarly defined. In general, cases with no automation present will feature $\underline{l}(R) = \bar{l}(R) = 0$. Automation occurs whenever $\underline{l}(R) > \bar{l}(R)$ or $\underline{l}(R) = \bar{l}(R) > 0$. An immediate implication of Proposition 6 is that for any $R' > R$ we have $\bar{l}(R') \geq \bar{l}(R)$ and $\underline{l}(R') \leq \underline{l}(R)$. We sharpen the characterization of the region of automation with the following result:

Proposition 7. *Let the assumptions of Proposition 6 hold. Consider two output levels R, R' with $R' > R$. Consider a step of length $l_i = \bar{l}(R)$ or of length $l_i = \underline{l}(R)$. We have two cases of interests:*

(i) Suppose $r_m(l_i, R) = r_m^*(l_i)$, then in the R' scenario the step is not automated: $o_i(R') = h$;

(ii) Suppose $r_m(l_i, R) < r_m^*(l_i)$, then in the R' scenario the step is automated: $o_i(R') = m$.

Proof. We focus on the case $l_i = \bar{l}(R)$. The case for $l_i = \underline{l}(R)$ follows in a similar manner. *(i)* Suppose not. We then have $o_i(R') = m$. In addition, $r_m(l_i, R') = r_m(l_i, R) = r_m^*(l_i)$, so that

$$\frac{w\left(D\left(c(l_i|\rho_m), r_m|m\right)\right)}{r_m} < \min_{r \leq R'\bar{r}_h} \left\{ \frac{w\left(D\left(c(l_i|\rho_h), r|h\right)\right)}{r} \right\},$$

since the $l_i = \bar{l}(R)$, the step was not automated so we also have that

$$\frac{w\left(D\left(c(l_i|\rho_m), r_m|m\right)\right)}{r_m} \geq \frac{w\left(D\left(c(l_i|\rho_h), r_h(l_i, R)|h\right)\right)}{r_h(l_i, R)},$$

Since the choice of $r_h(l_i, R)$ is available for the scenario with $R' > R$ the two above equations lead to a contradiction.

(ii) Suppose not. We then have:

$$\frac{w\left(D\left(c(l_i|\rho_m), r_m(l_i, R')|m\right)\right)}{r_m(l_i, R')} \geq \frac{w\left(D\left(c(l_i|\rho_h), r_h(l_i, R')|h\right)\right)}{r_h(l_i, R')}.$$

As a fist step we show that $r_h(l_i, R) > r_m(l_i, R)$. Since $l_i = \bar{l}(R)$ we also have:

$$\frac{w\left(D\left(c(l_i|\rho_m), r_m(l_i, R)|m\right)\right)}{r_m(l_i, R)} = \frac{w\left(D\left(c(l_i|\rho_h), r_h(l_i, R)|h\right)\right)}{r_h(l_i, R)}. \tag{28}$$

Since $r_m(l_i, R) < r^*(l_i)$ it also follows that

$$\frac{w\left(D\left(c(l_i|\rho_m), r_m(l_i, R')|m\right)\right)}{r_m(l_i, R')} < \frac{w\left(D\left(c(l_i|\rho_m), r_m(l_i, R)|m\right)\right)}{r_m(l_i, R)}.$$

If $r_h(l_i, R) = r_h(l_i, R')$ the previous three equations lead to a contradiction. It then follows that $r_h(l_i, R) < r_h(l_i, R')$ and hence $r_h(l_i, R) < r_h^*(R)$. This implies that both human and machine operators assigned to step l_i with output R are constrained: $r_m(l_i, R) = R\bar{r}_m$ and $r_h(l_i, R) = R\bar{r}_h$. From Assumption 5 it then follows that $r_h(l_i, R) > r_m(l_i, R)$. From (28) since $r_h(l_i, R) > r_m(l_i, R)$ we have that $D\left(c(l_i|\rho_h), r_h(l_i, R)|h\right) > D\left(c(l_i|\rho_m), r_m(l_i, R)|m\right)$. From Assumption 2, it follows that an increase in r will raise difficulty for the human performer more than for the machine performer. Consider now a level of output $R'' = R + \varepsilon$ with $\varepsilon > 0$ and small. Since the increase in difficulty is higher for the human performer and the gain of an increase in r is smaller for the human performer (recall that $r_h(l_i, R) > r_m(l_i, R)$) we then have that $o_i = m$ at R''. From Proposition 6 it also follows that $o_i = m$ at $R' > R''$, reaching a contradiction. □

4.4 Fragmentation Costs And Division of Tasks

In Section 2 we discussed how a significant source of technological change is the change in fragmentation costs. The goal of this section is to explore the relationship between changes in fragmentation costs and the implied changes in the division of tasks and changes in ability demand.

The environment allows for an arbitrarily complex pattern of fragmentation costs. To make progress, this section considers two benchmarks: a uniform fragmentation cost case, and then an arbitrary fragmentation cost case affected by a uniform change in fragmentation costs. As a first step we show that variation in fragmentation costs over production tasks is a necessary condition for wage inequality. Without variation in fragmentation costs, step length is uniform and hence so is the ability demand for performers.

Lemma 3. *Suppose that the Assumptions of Proposition 1 and Proposition 3 hold. Consider the case in which $f(\cdot,\cdot) = \bar{f}$. Then $l_i = \bar{l}$ and $a_i = \bar{a}$. for all $i = 1, \ldots, T$.*

Proof. Suppose not, then there exist two consecutive steps i, j such that without loss of generality $l_i > l_j$. Consider the alternative allocation with $\bar{l} = (l_i + l_j)/2$. For this allocation not to be optimal it must be the case that: $p(D\left(c(l_i|\rho_h), r_i|h\right), r_i, R|h) + p(D\left(c(l_j|\rho_h), r_j|h\right), r_j, R|h) \leq 2p(D(c(\bar{l}|\rho_h), \bar{r}|h), \bar{r}, R|h)$. The contradiction is then reached as in the proof of Proposition 1 exploiting sufficiently convex wages. The result for ability follows from Proposition 3. □

While constant fragmentation costs do not create heterogeneity in skill demand, the level of skill is impacted by the level of fragmentation costs. As a first step we show that a general reduction in fragmentation cost leads to an increase in the number of steps.

Lemma 4. *Suppose that the Assumptions of Proposition 3 hold. Suppose also that the function $w(\cdot)$ and $k(\cdot)$ are strictly convex. Consider an arbitrary profile for fragmentation costs f with an associated T thresholds. Consider an alternative profile for fragmentation costs f' so that $f'(t,\cdot) = f(t,\cdot) - \overline{f} > 0$ for all t. Let T' be the optimal number of thresholds under f'. We have that (i) $T' \geq T$. Suppose now that $w(\cdot)$ and $k(\cdot)$ are sufficiently convex, then: (ii) If f' sufficiently lower than f, then $T' > T$; (iii) For f' sufficiently low, then T' is arbitrarily large; (iv) let l_{min} (l'_{min}) be the length of the shortest step given T (T'), then for $\overline{f}$ sufficiently large, $l_{min} > l'_{min}$.*

Proof. We first establish that $w(\cdot)$ and $k(\cdot)$ convex implies $p(a,r)$ is convex in length. From Proposition 3, a_i minimizing $p(a,r)$ in (9) is increasing in l, so that by Assumption $w(a)$ or $k(a)$ is convex in length. Note also that r is weakly decreasing in l, so that for $g(R,r)$ weakly increasing and convex in r, so that $p(a,r)$ is also convex in length. This convexity holds even if the choice of performer changes in length: take any $l_i > l_j$ such that $h_i \neq h_j$, with $w(\cdot), k(\cdot)$ sufficiently convex by assumption. Then by cost-minimization, $p(a_j, r_j|h_i) \geq p(a_j, r_j|h_j)$, so that by $p(a,r|h+i)$ convex in l, h_j preserves convexity of p for steps of different length. The convexity of costs in length by Assumption and Proposition 3 holds even if there exists a constant $\underline{D}$ such that $a(l=0) > 0$. By the proposition, a is strictly increasing in l: for sufficiently convex w, any $a(l_i) - a(l_j) > 0$ ensures that $p(a,r)$ remains convex in length. For sufficiently convex $w(a)$ and $k(a)$ we can also ensure $p(a,r)$ sufficiently convex.

(i) Suppose not. Then $T' < T$. This implies that $C(R,T'|c') < C(R,T|c')$. Since the original allocation under cost c is feasible, it follows that:

$$C(R,T'|c') = \sum_{i=1}^{T'} p(a'_i, r'_i, R|o'_i) + \sum_{i=1}^{T'} f'(s'_i, o'_i) < \sum_{i=1}^{T} p(a_i, r_i, R|o_i) + \sum_{i=1}^{T} f'(s_i, o_i) = C(R,T|c) - T \cdot \overline{f}.$$

From the statement of the Lemma given the optimality of T we have:

$$C(R,T|c) = \sum_{i=1}^{T} p(a_i, r_i, R|o_i) + \sum_{i=1}^{T} c(s_i, o_i) \leq \sum_{i=1}^{T'} p(a'_i, r'_i, R|o'_i) + \sum_{i=1}^{T'} c(s'_i, o'_i) = C(R,T'|c') + T' \cdot \overline{f}.$$

Combining the two previous equations we get" $0 < \overline{f} \cdot (T' - T)$ reaching a contradiction if $T' < T$. The proof of *(ii)* and *(iii)* follows the proof of Proposition 1. *(iv)* First note that $l_{min} \leq \frac{\bar{v}}{T}$. If not $\sum_{i=1}^{T} l_i > \bar{v}$ reaching a contradiction. From *(iii)* for sufficiently high f' we have we have T' sufficiently greater than T to ensure $l_{min} > l'_{min}$. □

The previous Lemma enables us to determine the changing demand for ability as fragmentation costs change. As fragmentation costs decrease, an immediate implication of Lemma 4 and Proposition 3 is a decrease in the lowest ability demanded. The following corollary formalizes this statement.

Corollary 1. *Suppose the Assumptions of Lemma 4 hold. Consider an arbitrary profile for fragmentation costs f. Consider an alternative profile for fragmentation costs f' so that $f'(t,\cdot) = f(t,\cdot) - \overline{f}$ for all t. Let*

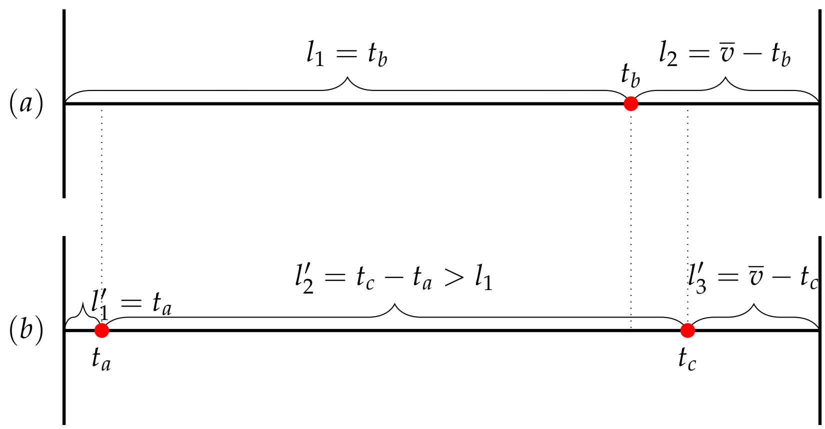


Figure 3: Maximum Step Length and T. (a) Case $T = 1$, (b) Case $T = 2$.

a_{min} (a'_{min}) be the lowest ability demanded under fragmentation cost f (f'). If f' sufficiently lower than f, then $a'_{min} < a_{min}$.

The last result in Lemma 4 shows how a sufficiently large reduction in fragmentation costs results in a decrease in the minimum step length. There is not an equivalent property for the maximum step length in a process. Indeed it is possible for the maximum step length to increase due to an increase in T, even if the total cost of production decreases. The following example makes this point.

Example 2. *In this example we show how it is possible for the longest step length to increase as the number of steps increases. Let fragmentation costs be arbitrarily high for all t, except for three points, t_a, t_b, t_c, with $f(t_a, \cdot) = f(t_c, \cdot)$ and let $f(t_b, \cdot) = f(t_a, \cdot) + d$, with $d > 0$. Let $t_a < t_c - t_b$ and $t_b > \frac{\bar{v}}{2}$. For $T = 1$ we have $s_1 = t_b$, the more centrally located cut. This is the case whenever the convexity of costs with respect to length dominates the higher fragmentation costs at t_b. For $T = 2$ we have $s_1 = t_a$ and $s_2 = t_c$. This occurs if the reduction in performer costs from placing step thresholds at t_a, t_b or t_b, t_c relative to t_a, t_c are less than a. Figure 3 summarizes the example when $T = 1$ and $T = 2$. The parametric scenario with $p(l) = l^{1.088}$, $t_a = .4, t_b = 7, t_c = 7.5, \bar{v} = 10, d = 0.05$ delivers the required properties for the example.*[16]

While the effect of a change in fragmentation costs on the maximum step length is indeterminate under arbitrary conditions, we can place additional structure on fragmentation costs to restrict the effects on maximum length. If there exists any interval of tasks $V = [t_i, t_j]$ such that $f(t)$ is arbitrarily high for $t \in V$, then the maximum step length will never be less than $t_j - t_i$. In this case we refer to V as a set of *lumpable-tasks*. It is natural to think of maximum step lengths being defined by regions of tasks which are indivisible or have arbitrarily high fragmentation costs, such as in highly controlled material deposition (e.g. in optoelectronics Combemale et al. 2020), continuous processing (e.g. in steel Goldin and Katz 1998), or indivisible loads in computing (e.g. Berenbrink et al. 2015). Under this additional structure, maximum step length is

[16] Formally, we require t_a, t_b, t_c be such that $p(t_b) + p(\bar{v} - t_b) < \min\{p(t_c) + p(\bar{v} - t_c) - a, p(t_a) + p(\bar{v} - t_a) - a\}$. Let t_a, t_c be such that $p(t_c - t_a) + p(\bar{v} - t_c) < p(t_c - t_b) + p(\bar{v} - t_b) + a$ and $p(t_c - t_a) + p(t_a) < p(t_b - t_a) + p(t_b) + a$.

insensitive to reductions in fragmentation cost, while the minimum step length is not. It is thus possible for technological changes to affect the upper and lower bounds of step complexity independently. These independent effects, as we show next, allow for technological changes that change the difference between the least and highest ability demand.

Changes in Issue Arrival Important historic technological changes have simultaneously affected the complexity and divisibility of processes.[17] Within the framework presented in this paper these technological changes can be described by a simultaneous change in fragmentation costs f and in the parameter governing the average number of issues λ.[18] In what follows, we consider the ability demand implications of a change in technology which increases issue arrival and sufficiently decreases fragmentation costs. We show that this change generates an increase in the upper bound of ability demanded and hence upward pressure on the highest wages in labor market equilibrium.

Corollary 2. *Suppose that the Assumptions of Proposition 3 hold. Suppose there exist a set of lumpable tasks $\widehat{V}$ of length $\widehat{l}$. Suppose also that under issue arrival λ the maximum step length is $\widehat{l}$. Consider an issue arrival $\lambda' > \lambda$. Let a_{max} (a'_{max}) be the lowest ability demanded under fragmentation cost λ (λ'). If the performer for the longest step remains the same, we then have $a'_{max} > a_{max}$.*

Proof. Since $\widehat{V}$ is lumpable, the maximum step length cannot be smaller than $\widehat{l}$ under any technological change. From the definition of complexity in (2) we observe that step length and issue arrival are perfect substitutes in their effect on complexity. The result then follows the proof of Proposition 3 substituting changes in l with changes in λ. □

The previous Corollary requires a constant performer type for the longest step. If the longest step is sufficiently long, then by Proposition 4 this Assumption is automatically satisfied, as human performers are assigned to this step before and after the change in issue arrival rate.

The previous Corollary together with Corollary 1 imply that in the present of technological change that simultaneously lowers fragmentation costs and raises issue arrival we will expect an increase of within plant inequality. Together Corollary 1 and 2 provide a theoretical basis to understand how within-firm inequality might increase or decrease given different types of technological change.

[17] For instance, the development of the assembly line in manufacturing permitted a much finer division of tasks but entailed a more complex overall process with greater logistical and managerial requirements (Hounshell, 1985; Chandler, 1990). The more recent phenomenon of design modularity in programming and other design allows for easier separation of work but increases system complexity (Baldwin and Clark, 2003). The inverse of this trade-off is also possible. In modern manufacturing, parts consolidation, when formerly discrete parts are fabricated as one piece, makes dividing tasks more costly but also reduces the number of issues that might arise in assembly (Selvaraj et al., 2009; Combemale et al., 2020).

[18] Technological change affecting multiple dimensions are also intuitive from an adoption perspective. For example, a firm will not adopt a technology increasing fragmentation costs or issue arrival without an opposing effect reducing costs, such as reduced fragmentation cost or fewer issues.

5 Empirical Analysis

In this section, we provide empirical counterparts to the theoretical results of the preceding section. First, increasing complexity of production requires performers with higher ability working at lower rates. Second, a reduction in fragmentation costs leads to an increase in the number of steps. Third, a reduction in fragmentation costs and an associated increase in issue arrival rates leads to an increase in the upper bound of ability demand, and a decline in the lower bound of ability demand. Fourth, automation substitutes for workers of middle ability at low volumes; this cone of automation widens as volume increases, until automation substitutes for all but high ability.

5.1 Data Sources

In this section we use three datasets. Each dataset used provides detailed information on production operations. The three datasets are the *Hand and Machine Labor Study* of 1898; data on optoelectronic semiconductor manufacturing taken from Combemale et al. (2020); and data on contemporary auto-body assembly taken from Fuchs et al. (2008).

The Hand and Machine Labor Study. The first dataset comes from the Hand and Machine Labor (HML) study (Wright, 1898).[19] The original data collection for this study was conducted by the Bureau of Labor Statistics between 1894 and 1898, with the goal of investigating the effect of the use of machines on labor. The study covers 672 products across the agricultural, manufacturing, mining and transportation sectors. Detailed descriptions are provided for all products, which vary from harvesting hay, to watch manufacturing, to road repair and cargo loading. Processes range from cases with a handful of steps (e.g. harvesting hay), to cases with hundreds of process steps (e.g. watch manufacturing). Every step of every production process is coded in the data in terms of its motive power (e.g. hand, steam, water). Every product recorded in the hand and machine labor study is produced by exactly two separate processes: a "hand" process (a process that is relatively more manual), and a "machine" process (a process that is relatively more mechanized). The two processes represent a change in process structure and performer type to produce the same good with identical characteristics.[20] The data characterizes each process step-by-step, analogously to the structure of steps in our model: for example, the hand process for producing hay consists of 1) mowing grass, 2) tending hay, 3) raking hay, 4) cocking hay, 5) hauling hay, 6) baling hay and 7) weighing hay. The data includes the occupations employed in each step, the number of employees for each occupation for the step, the task content of the step and the motive power used in the step (e.g. hand, or different types of machine power such as water, steam, or electricity). Wages and operations data consist of the time worked per step cycle, the output

[19] The dataset is also described in Atack, Margo, and Rhode (2019).

[20] Rare exceptions noted by the original authors of the HML study include slabs of granite of different final weight. These products are of the same composition, but different quantities, cut from similar raw materials (e.g. the walls of a quarry).

per cycle of a process step, the number of workers required per step and the number of workers required per workstation. Each process step has a detailed task description and coding to identify which step (or steps) in the hand process contains the same tasks as the machine process. For example, the machine process for making a sleigh (Product 183) includes steps coded 2 and 3 for sanding panels and setting up the sleigh body, while the hand process has a step for setting up and sanding the body, coded as (2,3) to indicate that it contains the same tasks. Refer to Appendix B for further details on this unique dataset.

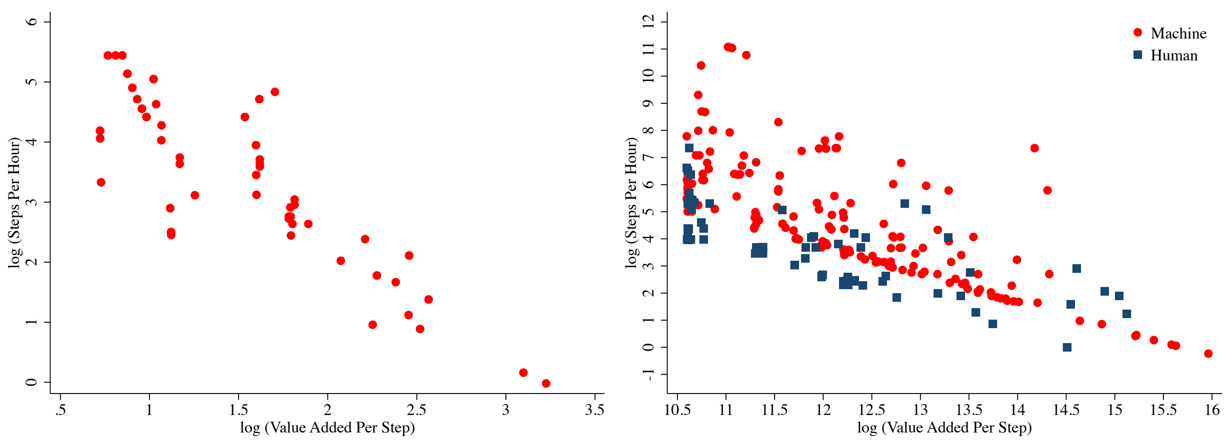
Direct Measurement of Production The remaining two datasets are direct measurement of plant-level production processes. This data is collected to identify the technical parameters of a highly detailed production model. These models, called Process Based Cost Models (PBCMs) in the industrial engineering and operations literature, are used across a variety of industries to inform engineering and production decisions. PBCMs describe the production of a single product in individual process steps and map characteristics of the product design (e.g. geometry) and process design (e.g. level of automation) to production outputs. This modeling approach provides the additional benefit of isolating the effects of technology changes at the level of individual process inputs, for example the effect of using a human or a machine to perform a specific production step on output. See Combemale, Whitefoot, Ales, and Fuchs (2020) for a detailed description of these models and the data collection process.

Optoelectronic Semiconductor Manufacturing The first production process is the production of optoelectronic semiconductor transceivers for communications. We use data from the fabrication of semiconductor components to their assembly into a final package. The optoelectronic semiconductor industry is a useful case study for the effects of technology change because optoelectronic transceivers have a common form factor and end-use, so that they are functionally homogeneous while varying significantly in their internal design and method of production (i.e. in terms of the technological parameters in our theory). The optoelectronic semiconductor manufacturing dataset was originally collected and presented in Combemale et al. (2020). This dataset allows us to compare step-level demand for worker skills (captured using the same methodology as the O*NET database) under different technological scenarios. These scenarios vary in the level of automation and the level of consolidation of product designs (increasing in the number of internal components which are jointly fabricated).

Automobile Body Fabrication and Assembly The final dataset is from automobile body fabrication and assembly. This dataset was originally collected and presented in Fuchs, Field, Roth, and Kirchain (2008). The data which we use in this paper characterize process flow and step-level process inputs for automobile body assembly. For each assembly process step, the data includes capital and labor inputs (demand, price) for each process cycle as well as operations parameters, specifically batch size and cycle time. The dataset also includes data for each step on the number of welding joins required for each part of the automobile body.

5.2 The Relationship Between Ability, Rate and Complexity

In this Section, we use the optoelectronic semiconductor and automobile body production data to provide an empirical analogue to Proposition 3, which relates rate (r) and ability (a) to step length (l). Following Remark 1, we use value added per step as a proxy for step length (this approximation relies on the second result in Proposition 3 relating ability and step length: longer steps have higher performer costs and hence higher value added). We calculate value added per step from the cost of labor and capital inputs to produce a unit of output from a step. We consider all human or machine inputs to execute a step.[21] The empirical results from both contexts, presented in Figure 4, are consistent with Proposition 3 by showing that rate is decreasing in step length. In the optoelectronic semiconductor context, the same wire-bonding machine takes longer to complete more complex configurations while preserving the same proportion of successful versus failed outputs. In the automobile body assembly context, more complex welding operations require more expensive machines (see Figure C.1 in Appendix) or require the same machines to operate more slowly (in the case of human operators more complex steps are often associated with more expensive tooling).



(a) Automotive

(b) Optoelectronics

Figure 4: Motivation for model ingredients. Data for (a) is from Fuchs et al. (2008). Data for (b) is from Combemale et al. (2020).

We next use the step-level skill measures from the optoelectronic semiconductor data to explore the relationship between a and l. Recall that Proposition 3 provides conditions in which a is strictly increasing in l. For this exercise we proxy step length l using the labor cost of producing a single unit of output from a process step.[22] In practice, production activity is shared in many steps between humans and machines, so to isolate the relationship between l and human ability demand we consider only steps in which human labor costs are at least 70 percent of value added

[21] Human costs are given by the compensation and tooling costs needed to execute those steps. Machine costs are given by the cost of the machine used, scaled by the time of use and the length of service life of the machine. For additional information refer to Combemale et al. (2020).

[22] Due to limitations of the data, we use a constant operator wage across all steps based on the average at each plant.

(the following results are robust to reducing minimum value added share to 60 percent).We compare across all steps in the dataset which have either a dexterity level of 1 (the lowest value) or a level of 5 (the highest value recorded in the data).[23] The distribution of labor costs associated with steps of high and low skill is consistent with Proposition 3 that highlights a negative relationship between ability and length. We have that average labor cost per unit for low-skill steps (19 observations) is \$0.19, while the average labor cost per unit for high-skill steps is \$0.52.

5.3 Fragmentation Costs and Division of Tasks

The time period covered by the HML dataset is characterized by a reduction of fragmentation costs.[24] This dataset thus offers a useful empirical counterpart to the results of Section 4.4. As a first step we look at how the historical general reduction in fragmentation costs leads to changes in the number of production steps. In the HML dataset we look at mappings between hand and machine process steps to capture intervals which are affected by an increase in the division of tasks (for detailed information on the processing of data, refer to Appendix B.2). We focus on the overall distribution of step lengths across all processes, because the HML dataset does not contain direct information on the fragmentation costs for each process. Results are in Figure 5. The figure displays a reduction in the number of processes that feature a small number of steps,

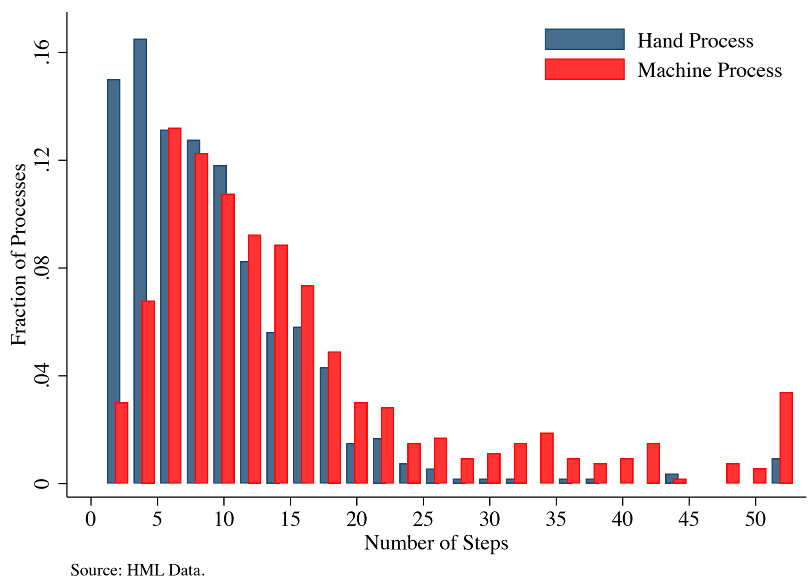


Figure 5: Fragmentation costs and step divisibility. HML data.

consistent with the increased division of tasks provided by Lemma 4. In the context of our theory

[23] As a check for robustness to skill type, we also performed a comparison between all steps whose maximum skill was 1 across the skills captured in the data (operations and control, dexterity, near vision) and steps whose maximum skill was 5. We found comparable results.

[24] For a historical account of the implications of technological change on manufacturing refer to Hounshell (1985), Chandler (1990).

(see Lemma 4) this phenomenon is rationalized by a reduction in fragmentation costs.

We next move to wages. Figure 6 is based on the HML data set. In this figure we restrict the analysis to steps with constant performer type (i.e. manual motive power in hand and machine processes). The Figure displays the moments of four distinct distributions. With the leftmost two box charts, we compare the distribution of relative wages for tasks in which changing from the hand to the machine Process does not incur changes in number of steps to perform those tasks (constant T). With the rightmost two box charts we look at the case of tasks for which changing from the Hand to the Machine Process leads to an increased number of steps (T increasing). In either case we compare the distribution of relative wages in the case of hand and machine processes.

In Figure 6 the behavior of wages and ability is consistent with Corollary 1. For the case of increasing T, we observe a decrease of the lowest wages. As wages are monotone in abilities this also suggests a decrease of ability demanded.[25] The Corollary emphasizes how in the presence of decreasing fragmentation costs we expect downward pressure in demand for the lowest skills. To confirm that the changes in wages are indeed driven by changes in the number of steps, in the left panel we also consider products for which the number of steps did not change as the process moved from hand to machine. In this case, the widening of the distribution of relative wages is much smaller than the case of increasing T. In Figure 6, for the case of increasing T, we

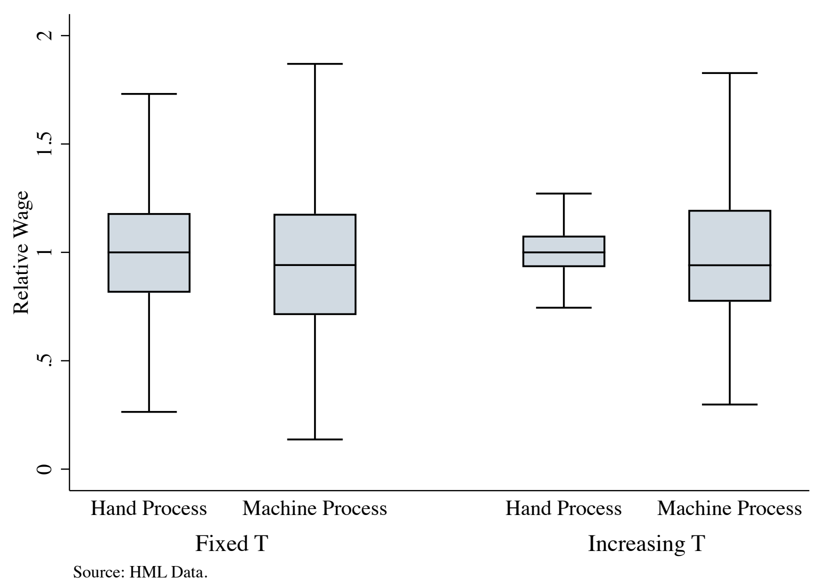
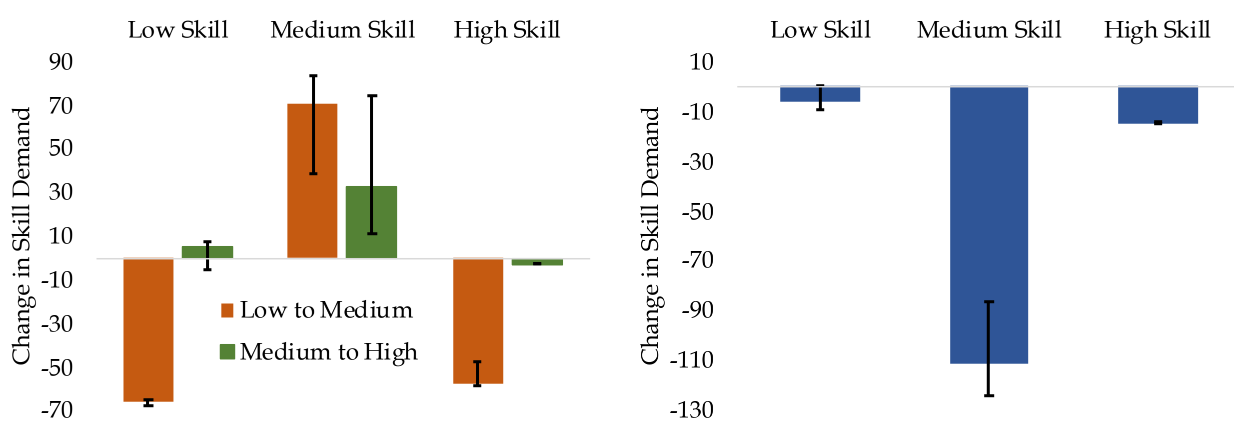


Figure 6: Fragmentation costs and wages. HML data.

also observe an increase in the relative wages at the top. Following the result in Corollary 2, this phenomenon can be rationalized by an increase in issue arrival.

[25]This pattern appears in direct industry observations. For example Womak et al. (1990) confirms the shortening of steps as the Automotive industry moved away from the hand process. The cycle time of workers between 1908 and 1913 decreased from 514 to 2.3 minutes. Similarly the training required for workers declined also to a few minutes.

The HML dataset has the advantage of covering a variety of different industries and products. A downside is the lack of precise controls on the nature of fragmentation cost and lack of precise measurements of ability levels. To overcome these limitations we next look at the data from the optoelectronic semiconductor and automobile assembly. This modern production data allows for direct observation of ability and precise control over the changing nature of the process as fragmentation costs and average number of issue change. The optoelectronic semiconductor production data features adoption of different levels of automation and consolidation. For all levels of automation and consolidation, the final products are functionally homogeneous and perfect substitutes on the market. Changing the level of consolidation of the design drives step consolidation: the more consolidated the design, the fewer the step thresholds (T). Consolidation of parts leads to an increase in fragmentation costs (f) but also a reduction in assembly requirements, captured in our theory by reduced issue arrival (λ). The case of consolidation allows us to look for an empirical analogue of Corollary 1 and Corollary 2 for constant performer type. Taking these two Corollaries together, we expect a convergence in ability demand (decline at the top and at the bottom), as fragmentation costs increase and issue arrival decrease (this is the opposite scenario as the one studied in the HML dataset). We use the skill-ratings collected for each step by Combemale et al. (2020) as a measure of a. Holding performer type constant across levels of consolidation, Figure 7a shows the effects of two changes in consolidation (from low to medium consolidation and then from medium to high consolidation) on the distribution of skill demand. We see that with consolidation skill demand converges toward middle skills.



(a) Parts Consolidation.

(b) Automation

Figure 7: Impact of technological change on skill demand. In (a) different color bars denote different levels of parts consolidation. Data from Combemale et al. (2020).

5.4 Automation

The next set of empirical results relates to which steps in a production process are most likely to be automated. The results in Section 4.3 provide guidance on what steps are more likely to be automated when considering steps of different length or production processes with different levels of output. Together, the results of Section 4.3 describe what we refer to as a *cone of automation* where automation is more likely for higher level of output and for middle length steps (see Figure 2). In what follows we first look at the HML data and then at the optoelectronic semiconductor data to find evidence for this pattern of automation.

We begin with the HML dataset. Ideally, to look at an empirical analogue for Figure 2, we need precise measurement of step length and observations of the production process at different output levels. The HML dataset provides a proxy of step length by providing wage data; unfortunately a process is observed at only one output level. Nevertheless, the HML dataset can be used for an empirical counterpart of Figure 2. The key insight, described below, is to use the capacity utilization of each step as a proxy for the overall output level. The variety of processes observed across different products (correctly scaled) then provides variation in capacity utilization across steps. Before proceeding note that Figure 2 describes strict upper and lower bounds on automation for a given set of structural parameters. In the approach that follows, we compare different products in the HML dataset. Intuitively, the different products are heterogeneous in the production structural parameters (for example, they might differ in ρ or σ). Given this unobserved heterogeneity we expect to observe a probability of automation that varies as we vary wages and capacity utilization as opposed to a strict demarcation.

In the model, changes in R impact the firm choices by affecting the cost of performers. This occurs through the function $g(\cdot,\cdot)$. As it can be observed in Assumption 3, the role of R is symmetric to the role of the minimum divisibility threshold of the performer $\bar{r}$. So that an impact of an increase in R can be achieved by an increase in $\bar{r}$. With this logic, the results of Section 4.3 and the pattern in Figure 2 can be recast in terms of $\bar{r}$ as opposed to R. This is helpful since the HML data, while not providing observation for different R allows us to recover the $\bar{r}$ for each process. We can then look at each manual step and determine how close the rate of the operator is to the minimum divisibility threshold (a sign of high capacity utilization). Using the same logic as Section 4.3, steps with high capacity utilization of human performers are more likely to be automated. We next briefly describe how to $\bar{r}$ is identified in the HML data (see Appendix B.2 for additional details). The dataset provides information on the number of workers involved in a step and the amount of time the step requires in order to be completed. For each process, following Hopp and Spearman (2011), we identify the bottleneck in production by looking at the step that requires the longest time to be completed. We determine the fractional utilization of a step by comparing its completion time to the completion time of the bottleneck. For example if a bottleneck requires 10 hours to be completed and a preceding step requires 1 hour to be completed, the fractional utilization of the preceding step is 1/10. Finally using the information on the number of performers on a given step we recover the fractional utilization of performers in a step. In the previous example, if the step has two workers assigned to it, it implies that the fractional utilization of performers in the step is 1/5 of a worker. This fractional performer utilization rate can be compared across steps and across processes and is used as one of the two

Wage Bin	1	2	3	4	5	6	7	8
8	17%	29%	17%	25%	30%	33%	44%	40%
7	25%	21%	35%	14%	47%	58%	63%	74%
6	25%	52%	50%	40%	52%	48%	58%	93%
5	17%	N/A	30%	57%	63%	40%	78%	91%
4	35%	60%	56%	70%	62%	68%	76%	91%
3	38%	53%	59%	44%	67%	68%	91%	91%
2	39%	42%	40%	35%	72%	71%	76%	77%
1	30%	41%	40%	41%	29%	65%	65%	64%
	Step Utilization Bin							

Figure 8: Patterns of automation over wage and utilization bins. Numbers in each cell denote the percentage of steps automated in each step. HML data.

key drivers for automation in Figure 8.

The second dimension driving automation in Figure 2 is step length. The connection of this dimension to the data is more straightforward. Proposition 3 establishes a direct relation between step length and ability. As ability is not observed in the HML dataset we proxy this characteristic using worker wages (in the optoelectronic semiconductor data below we instead directly observe ability). For each production process, wages for each step are normalized by the average wage observed in that process.

In the HML dataset, for each product we consider pairs of steps with identical task content between the hand and machine processes. We select steps from hand processes that were performed manually. We then measure whether a step has been automated in the machine process using a binary indicator of a change in motive power.[26] Figure 8 displays the results. In the Figure, each cell is ordered in terms of percentile of performer utilization and wage of the performer. The number in each cell denotes the percentage of steps in each range of utilization and wage that is automated. As expected from Figure 2, the pattern that emerges displays characteristics of a *cone of automation*: automation occurs more often at middle wage steps, and the range of middle wage steps that are likely to be automated becomes progressively larger for higher utilization steps. Intuitively, the most automated steps in the HML data are the ones in which a large fraction of worker time is devoted to a step thus allowing a machine in that step to be less rate-constrained. Additionally automation is more likely when wages are not too high or too low (so that machines are not executing a too complex step and as before are not rate-constrained).

We next turn to optoelectronic semiconductor data. In this dataset we observe different plant scenarios with different levels of automation.[27] This level of detail allows us to precisely determine if a step has been automated or not. In addition, the available data allows us to determine

[26]We do not observe any instances of a hand step shifting to a less mechanized form of motive power in the equivalent step in the machine process. For additional details refer to Appendix B.2 and B.3.

[27]The change in level of automation is characterized using a taxonomy of automation, For additional details refer to Combemale et al. (2020).

the ability of each operator without relying on data. The data provides information on ability levels as defined in the O*NET database.[28] In Figure 7b we display the results as we move from a low to medium level of automation. The data displayed is for a single output level. This Figure can then be considered as a vertical slice of Figure 2 in a region where automation occurs. The vertical axis denotes the number of displaced workers being automated at a given skill level. As anticipated by our theory, the impact of automation is more evident for middle-skill workers.

6 Concluding Remarks

In this paper, we provide a general theory relating technology change and labor demand. We emphasize three dimensions of the problem of the firm which are affected by technology change: the ease of fragmenting the production process into smaller steps (with associated changes in process complexity); the costs of allocating the same performer (human or machine) to multiple steps; and the trade-off between step complexity and rate of completion, where humans tend to solve more complex steps than machines but tend to have a slower completion rate.

We find that implications of the theory are supported by empirical evidence across a range of technologies and industry contexts from the late 19^{th} century and contemporary production. Human performers are favored against machines in high complexity steps, or at low volumes, so that automation has a polarizing effect on skill demand at low volumes and an upskilling effect at higher volumes. Technology changes that reduce fragmentation costs and increase process complexity can increase the spread of labor abilities demanded. Such technological changes put upward pressure on wage inequality. Conversely, we also find that technologies that increase fragmentation costs and lower complexity reduce the spread of labor abilities demanded, putting downward pressure on wage inequality.

The theory offers multiple broad insights for technology change and the division of labor. First, the division of tasks is the origin of heterogeneous ability demand. Heterogeneous ability demand will not occur unless some tasks are more costly to divide than others, and technological change can affect wage inequality by altering these fragmentation costs. Second, as machines become more divisible (e.g. through cloud computing), the effect of automation becomes less polarizing and more upskilling. Third, the singularity point where all human labor is replaced by machines occurs when machines are general enough and cheap enough such that they no longer have a relative disadvantage in solving either simple or complex problems compared to humans.

The theory also offers a unified explanation of technological change, capable of rationalizing a large number of empirical regularities observed in the data. In Table 1 we provide a mapping between major technological trends of the last two centuries and deep parameters and functional forms of our model. Based on the theory, a taxonomy of technology change can be developed. A technology change may be described in terms of its effects on process complexity (λ), task

[28] A skill of 1 is rated low, a skill of 5 or greater (levels 6 and 7 were not observed) high, and 2-4 medium. As shown in Combemale et al. (2020), this result is robust across different types of skills and without aggregation of skill rankings.

separability (f) and performer characteristics such as divisibility (g), sensitivity of performers to rate (related to σ) and generality (related to ρ). The classification of technology changes into their effects on these parameters suggests the resulting impacts of the technology on labor demand.

Table 1: General theory applied to a variety of technology changes.

Technology Change	Period	Theory Interpretation	Labor Impact
Mechanization: Substitution of human performers by machines	1870s-1890s	Machine performers repeat tasks faster than humans but unable to perform highly varied work: $\rho^m < \rho^h$ and $\sigma^m < \sigma^h$. Machines less divisible than humans, $\bar{r}_m < \bar{r}_h$.	Human ability demand polarized. Empirically: growth of higher skill professional jobs (Chandler, 1990), more demand for unskilled labor (Atack et al., 2019)
Interchangeable Parts and Assembly Line: Increased standardization of parts and process to facilitate transfer of work and minimize refitting requirements	1870s-1910s	Increased process complexity, leading to $\lambda \uparrow$, but facilitation of transfer and reduced post-processing of parts driving $f \downarrow$	Upper bound of human ability demand increases, lower bound of demand decreases. Empirically: creation of new managerial jobs and of very simple production jobs. (Hounshell, 1985; Womak et al., 1990)
Consolidation of Parts: Formerly discrete parts fabricated as one	1970s-2010s	Joint fabrication of parts makes some fabrication tasks indivisible, driving $f \uparrow$, allows simpler design and reduced assembly, driving $\lambda \downarrow$	Upper bound of human ability demand decreases, lower bound increases. Empirically: convergence of skill demand from low and high to middle, reduced division of production (Combemale et al., 2020)
Automation and Computerization: Substitution of human labor by computer and machine performers	1960s-2010s	Machine performers able to repeat tasks faster than humans but unable to perform highly varied work: $\rho^m < \rho^h$ and $\sigma^m < \sigma^h$. Compared to mechanization, performers are more general ($\rho \uparrow$), intense ($\sigma \downarrow$) and divisible ($\bar{r} \uparrow$)	Polarization of worker ability demand at low volumes, shifting to high skill at high volumes. Empirically: up-skilling of skill demand (especially in manufacturing), aggregate polarization in conjunction with lower automation in services (Goos et al., 2019; Willcocks and Lacity, 2016)

While the model is rich enough to model the skill demand implications of an unprecedented variety of technology changes, we also anticipate a number of extensions. A natural extension is to relax the assumption in the model that firms set their ability demand to ensure that a step is completed in expectation. This extension would allow firms to choose ability greater or less than difficulty at the cost of higher or lower yield rates. This extension could be important because, in certain empirical cases, high costs of failing to solve issues in a specific step would warrant higher demand for operator ability so that failure is less frequent. An additional extension could be imposing additional costs associated with the reorganization of a process. This extension would allow us to distinguish the effects of technology change with and without process reorganization, and could be extended to understand the incentives for different strategies of technology adoption.

References

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics 4*, 1043–1171.

Acemoglu, D. and P. Restrepo (2018a). Low-skill and high-skill automation. *Journal of Human Capital 12*(2), 204–232.

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

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

Ales, L., M. Kurnaz, and C. Sleet (2015). Technical change, wage inequality, and taxes. *American Economic Review 105*(10), 3061–3101.

Atack, J., R. A. Margo, and P. W. Rhode (2019). Automation of manufacturing in the late nineteenth century: The hand and machine labor study. *Journal of Economic Perspectives 33*(2), 51–70.

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

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

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.

Baldwin, C. Y. (2008). Where do transactions come from? Modularity, transactions, and the boundaries of firms. *Industrial and corporate change 17*(1), 155–195.

Baldwin, C. Y. and K. B. Clark (2003). Managing in an age of modularity. *Managing in the modular age: Architectures, networks, and organizations 149*, 84–93.

Becker, G. S. and K. M. Murphy (1992). The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics 107*(4), 1137–1160.

Berenbrink, P., C. Cooper, T. Friedetzky, T. Friedrich, and T. Sauerwald (2015). Randomized diffusion for indivisible loads. *Journal of Computer and System Sciences 81*(1), 159–185.

Biau, G. and D. M. Mason (2015). High-dimensional p-norms. In *Mathematical Statistics and Limit Theorems*, pp. 21–40. Springer.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics 117*(1), 339–376.

Brynjolfsson, E., T. Mitchell, and D. Rock (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings 108*, 43–47.

Cai, J. (2014). *Cramér–Lundberg Asymptotics*. Wiley Online Library.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of management review 13*(1), 40–52.

Caroli, E. and J. Van Reenen (2001). Skill-biased organizational change? Evidence from a panel of British and French establishments. *The Quarterly Journal of Economics 116*(4), 1449–1492.

Caselli, F. (1999). Technological revolutions. *American economic review 89*(1), 78–102.

Chandler, A. D. (1990). *Scale and Scope: The Dynamics of Industrial Capitalism*. Harvard University Press.

Combemale, C., K. S. Whitefoot, L. Ales, and E. R. Fuchs (2020). Not all technological change is equal: Disentangling labor demand effects of automation and parts consolidation. Forthcoming at Industrial and Corporate Change.

Costinot, A. and J. Vogel (2010). Matching and inequality in the world economy. *Journal of Political Economy 118*(4), 747–786.

Dinlersoz, E. and Z. Wolf (2018). Automation, labor share, and productivity: Plant-level evidence from US manufacturing. CES Working Paper 18-39.

Eden, M. and P. Gaggl (2019). Capital composition and the declining labor share. CESifo Working Paper.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology 47*(6), 381.

Frohm, J., V. Lindström, J. Stahre, and M. Winroth (2008). Levels of automation in manufacturing. *Ergonomia-an International journal of ergonomics and human factors 30*(3), 1–28.

Fuchs, E. R., F. R. Field, R. Roth, and R. E. Kirchain (2008). Strategic materials selection in the automobile body: Economic opportunities for polymer composite design. *Composites science and technology 68*(9), 1989–2002.

Fuchs, E. R., F. R. Field, R. Roth, and R. E. Kirchain (2011). Plastic cars in China? The significance of production location over markets for technology competitiveness in the United States versus the People's Republic of China. *International Journal of Production Economics 132*(1), 79–92.

Garicano, L. and E. Rossi-Hansberg (2006). Organization and inequality in a knowledge economy. *The Quarterly journal of economics 121*(4), 1383–1435.

Goldin, C. and L. F. Katz (1998). The origins of technology-skill complementarity. *The Quarterly journal of economics 113*(3), 693–732.

Goos, M., A. Manning, and A. Salomons (2009). Job polarization in europe. *American economic review 99*(2), 58–63.

Goos, M., E. Rademakers, A. Salomons, and M. Vandeweyer (2019). Job polarization: Its history, an intuitive framework and some empirical evidence.

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

Haanwinckel, D. (2020). Supply, demand, institutions, and firms: A theory of labor market sorting and the wage distribution. Working Paper.

Hopp, W. J. and M. L. Spearman (2011). *Factory physics*. Waveland Press.

Hounshell, D. A. (1985). *From the American System to Mass Production, 1800-1932: The Development of Manufacturing Technology in the United States*. JHU Press.

Kirmaci, U., M. Klaričić Bakula, M. Özdemir, and J. Pečarić (2008). On some inequalities for p-norms. *Journal of inequalities in pure and applied mathematics 9*(1), 27–1.

Korsah, G. A., A. Stentz, and M. B. Dias (2013). A comprehensive taxonomy for multi-robot task allocation. *The International Journal of Robotics Research 32*(12), 1495–1512.

Laureijs, R. E., E. R. Fuchs, and K. S. Whitefoot (2019). Is more less? Benefits and costs of high-variety production in nonassembled manufacturing. *Journal of Mechanical Design 141*(6), 1–30.

Lin, H.-I. and C. Lee (2013). Measurement of the robot motor capability of a robot motor system: A Fitts-law-inspired approach. *Sensors 13*(7), 8412–8430.

Lindenlaub, I. (2017, 01). Sorting Multidimensional Types: Theory and Application. *The Review of Economic Studies 84*(2), 718–789.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed–accuracy trade-off in skilled behavior. *Psychological Review 89*(5), 483.

Ocampo, S. (2018). A task-based theory of occupations with multidimensional heterogeneity. Working Paper.

Rosen, S. (1978). Substitution and division of labour. *Economica 45*(179), 235–250.

Schmitz, J. A. and A. Teixeira (2008). Privatization's impact on private productivity: The case of Brazilian iron ore. *Review of Economic Dynamics 11*(4), 745–760.

Selvaraj, P., P. Radhakrishnan, and M. Adithan (2009). An integrated approach to design for manufacturing and assembly based on reduction of product development time and cost. *The International Journal of Advanced Manufacturing Technology 42*(1-2), 13–29.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal 27*(3), 379–423.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. Printed for W. Strahan; and T. Cadell.

Staats, B. R. and F. Gino (2012). Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management science 58*(6), 1141–1159.

Stigler, G. J. (1951). The division of labor is limited by the extent of the market. *Journal of political economy 59*(3), 185–193.

Vicentini, F. (2021). Collaborative robotics: a survey. *Journal of Mechanical Design 143*(4), 040802: 1–20.

Welford, A. T. (1981). Signal, noise, performance, and age. *Human Factors 23*(1), 97–109.

Wickens, C. D., J. G. Hollands, S. Banbury, and R. Parasuraman (2015). *Engineering psychology and human performance*. Psychology Press.

Willcocks, L. P. and M. Lacity (2016). *Service automation robots and the future of work*. SB Publishing.

Womak, J., D. T. Jones, and D. Roos (1990). *The machine that changed the world*. New York: Rawson Associates.

Wright, C. D. (1898). *Hand and machine labor*. US Government Printing Office.

Yang, X. and S. Ng (1998). Specialization and division of labour: A survey. In *Increasing returns and economic analysis*, pp. 3–63. Springer.

Appendix

A Proofs of Section 3

Proof of Lemma 1

Proof. For any realization of $\{X_i\}_{i=1}^n$, from Lemma 2.1 in Kirmaci et al. (2008) we have that:

$$\left(\sum_{j=1}^{n}(X_j)^{\rho_h}\right)^{\frac{1}{\rho_h}} < \left(\sum_{j=1}^{n}(X_j)^{\rho_m}\right)^{\frac{1}{\rho_m}},$$

the result then follows immediately. $\square$

Proof of Lemma 2

Proof. We have that for all n

$$\frac{dP_n(l)}{dl} = P_n(l)\left(\frac{n}{l} - \lambda\right).$$

Since $\mathbf{X}(n|\rho)$ is strictly increasing in n, we have that:

$$\frac{dc(l)}{dl} = \sum_{n=1}^{\infty} P_n(l)\left(\frac{n}{l} - \lambda\right)\mathbf{X}(n|\rho) > \sum_{n=1}^{\infty} P_n(l)\left(\frac{n}{l} - \lambda\right)\mathbf{X}(1|\rho),$$

so that:

$$\frac{dc(l)}{dl} > \mathbf{X}(1|\rho)\left[\frac{1}{l}\sum_{n=0}^{\infty} nP_n(l) - \lambda(1 - P_0(l))\right] = \mathbf{X}(1|\rho)\lambda P_0(l) > 0.$$

We next show concavity. Since $\mathbf{X}(0|\rho) = 0$ and since $\frac{n}{l}P_n(l) = \lambda P_{n-1}(l)$, we have that

$$\frac{dc(l)}{dl} = \lambda\sum_{n=0}^{\infty} P_n(l)(\mathbf{X}(n+1|\rho) - \mathbf{X}(n|\rho)) > 0,$$

so that:

$$\frac{d^2c(l)}{dl^2} = \lambda\sum_{n=0}^{\infty} P_n(l)\left(\frac{n}{l} - \lambda\right)(\mathbf{X}(n+1|\rho) - \mathbf{X}(n|\rho)),$$

so that:

$$\frac{d^2c(l)}{dl^2} = \lambda^2\sum_{n=0}^{\infty} P_{n-1}(l)(\mathbf{X}(n+1|\rho) - \mathbf{X}(n|\rho)) - \lambda^2\sum_{n=0}^{\infty} P_n(l)(\mathbf{X}(n+1|\rho) - \mathbf{X}(n|\rho)) =$$
$$< \lambda^2\sum_{n=1}^{\infty} P_n(l)\left[(\mathbf{X}(n+2|\rho) - \mathbf{X}(n+1|\rho)) - (\mathbf{X}(n+1|\rho) - \mathbf{X}(n|\rho))\right] < 0.$$

Where the first inequality holds since $\lambda^2 P_0(l)(\mathbf{X}(1|\rho) - \mathbf{X}(0|\rho)) > 0$ and the second inequality holds since $\mathbf{X}(n+1|\rho) - \mathbf{X}(n|\rho)$ is decreasing in n. $\square$

B Hand and Machine Labor Data

B.1 Overview

For each product, the Hand and Machine Labor (HML) dataset includes general information on the process as well as detailed information on the steps required to create the product either using different methods: either by hand or using a machine. Products vary greatly in the complexity of production: the observed number of steps range from one to over two hundred and fifty. The data used in this paper comprise 247,482 step-level entries and 11,862 process-level entries. The HML data is publicly available in a non-digitized form. The entire dataset was digitized from scanned and physical copies of the data by undergraduate students at Carnegie Mellon University between 2019 and 2021.[1]

The type and definition of variables present in the HML dataset are described below. Table B.1 describes process-level variables, which apply across all steps and methods. For ease of comparison between processes of each method, the dataset reports observed production volumes for each process. The dataset also reports the input requirements to meet a "conformed"volume which is consistent across the hand and machine methods. Table B.2 describes variables which are reported for each step.

Table B.1: Process variables in HML data.

Variable Name	**Definition**	**Example**
Unit	Product name	Potatoes
Unit Volume	Volume of product captured for each full cycle of process	880 bushels
Conformed Volume	Volume of product per cycle used in presentation of step-level data	220 bushels
Method	Level of process mechanization	Hand/Machine
Total Employment	Number of people employed in process	4 people
Total Animals	Number and type of animals used in process	2 horses
Time Worked	The number of hours worked per day	10 hours
Year	Date of production process	1893
Unit Characteristics	Additional product details	From grafts

[1] Approximately 90% of products in the Hand and Machine Labor Study have been digitized so far.

Table B.2: Step variables in HML data.

Variable Name	**Definition**	**Example**
Operation Number	Identifying code for the set of tasks in a process step	{2, 3}
Work Done	Description of the activities performed in a step	Planting Seed
Machine, Implement or Tool Used	Description of primary equipment used to complete step	Steam shovel
Motive Power	Source of power for operations described	Steam; Horse
Persons Necessary on One Machine	Number of workers required per machine or station	2 workers
Animals Necessary on One Machine	Number of animals required per machine or station (type recorded in motive power)	2 horses
Number of Workers	Number of workers required in a process step across all stations	4 workers
Sex	Sex of workers	M, F
Occupation	Occupational title of workers	Laborer
Age	Age (or age range) of workers	21-30
Time Worked	Total person-hours and minutes to complete step	1hr 15m
Animal Time Worked	Total animal-hours and minutes to complete step	2hr 30m
Worker Pay Rate	Rate of pay (nominal dollars) for payment period	$1.00
Animal Pay Rate	Cost of animal (nominal dollars) for payment period	$0.375
Worker Pay Period	Payment cycle for workers	1 Day
Animal Pay Period	Cost cycle for animals	1 Day
Labor Cost	Total labor cost of producing conformed volume	$.125
Animal Cost	Total animal cost of producing conformed volume	$.0938

B.2 Mapping Hand and Machine Processes

We next describe the steps taken to map the data to the model. In the original data, entries concerning animal labor in production are given a distinct line with otherwise identical step information (tools, task content). Since there are never animals used in production without workers, we condense animal information into the same step as the human workers that manage them. Some steps also include workers with multiple occupational titles. When this occurs the dataset provides separate entries in the data. When mapping the task content between hand and machine methods, distinct occupations are kept as separate steps with the same task content. Any step containing multiple occupations (7.6 percent of steps observed) is excluded from our analysis of step automation or changes in the division of tasks among steps, because the division of tasks within occupations within a single step is not specified (and to avoid double-counting steps).

For all products, we build a mapping between hand and machine processes. We index the tasks in hand and machine processes as $\mathcal{V}^H$ and $\mathcal{V}^M$ respectively. In terms of notation, H, M indicate either hand or machine process-types. Every step i contains a set $\mathcal{S}_i$ of tasks. Note that it is possible for two steps $i \neq j$ to exist such that $\mathcal{S}_i^M \cap \mathcal{S}_j^M \neq \varnothing$: for example, steps with content 1a and 1b in Hand are identical in task content to step 1 in Machine, and to each other. Any given step belongs to exactly one of the following six possible cases:

1. **1 to 1:** Steps i^H and j^M belong to this case if they have the same task content and do not share task content with any other steps: $\mathcal{S}_i^H = \mathcal{S}_j^M$. For any $n \neq i$ then $\mathcal{S}_n^H \cap \mathcal{S}_j^M = \varnothing$, and for any $m \neq j$ then $\mathcal{S}_i^H \cap \mathcal{S}_n^M = \varnothing$. A 1 to 1 mapping is useful when analyzing a change in performer type or performer characteristics, independently of changes in the division of production.

2. **1 to 0:** a step i^H belongs to this case if $S_i^H \cap \mathcal{V}^M = \varnothing$. These steps capture activities that are no longer performed in the machine case (e.g. post-processing work made unnecessary by process improvement).

3. **0 to 1:** a step i^H belongs to this case if $S_i^M \cap \mathcal{V}^H = \varnothing$. These steps represent activities which are new to a process (e.g. firing a boiler, which would be unnecessary in a hand process without a steam engine).

4. **1 to N:** step i^H belongs to this case if: *(a)* $\mathcal{S}_i^H \subset \mathcal{V}^M$ (all of its tasks are contained in the machine process), *(b)* $\exists m \neq n$ such that $\mathcal{S}_n^M, \mathcal{S}_m^M \subset \mathcal{S}_i^H$ (tasks in the hand step are contained in more than one machine step), and *(c)* $\forall j$ such that $\mathcal{S}_j^M \cap \mathcal{S}_i^H \neq \varnothing$ we have $\mathcal{S}_j^M \cap (\mathcal{V}^M \setminus \mathcal{S}_i^H = \varnothing)$ (no machine step with a task set intersecting the hand step contains tasks that are contained in any other hand step:[2]) Step j^M belongs to this case if $\mathcal{S}_j^M \subset \mathcal{S}_i^H$ such that i^H satisfies the above conditions. This case allows us to capture an increase in the division of tasks.

5. **M to 1:** a step j^M belongs to this case if: *(a)* $\mathcal{S}_j^M \subset \mathcal{V}^H$, *(b)* $\exists m \neq n$ such that $\mathcal{S}_n^H, \mathcal{S}_m^H \subset \mathcal{S}_j^M$ and *(c)* for any i such that $\mathcal{S}_i^H \cap \mathcal{S}_j^M \neq \varnothing$, $\mathcal{S}_i^H \cap (\mathcal{V}^m \setminus \ \mathcal{S}_j^M) = \varnothing$. Step i^H belongs to this case if $\mathcal{S}_i^H \subset \mathcal{S}_j^M$ such that j^M satisfies the above conditions. This case allows us to capture a decrease in the division of tasks.

[2] Including tasks which occur in both step i^H and another hand step.

6. **M to N:** any remaining step not included above belongs to this case.

Table B.3 reports the number and share of process steps for each method which belong to each of the six cases described above. We see that 78.6% of Hand steps and 83.4% of Machine steps belong to mappings which can be interpreted as changes in T for fixed $\mathcal{V}$, allowing them to be used to explore technological cases which vary or hold constant the division of tasks.

Table B.3: Mapping between steps of different methods recovered from HML data.

Process Mapping	**Hand Steps**	**Share of Hand**	**Machine Steps**	**Machine Share**
0 to 1	0	0	2948	.333
1 to 0	204	.042	0	0
1 to 1	2375	.484	2375	.269
1 to N	639	.130	1921	.217
M to 1	639	.130	131	.015
M to N	1039	.212	1469	.166
Missing Alternate	9	.02	0	0
Total	4896		8844	

The "*Missing Alternate*"row indicates steps from processes which do not have a corresponding process of the opposite method: in our data, one hand process had a counterpart machine process for which the authors of the Hand and Machine study could not compare task content and thus could not encode operation numbers.

B.3 Measuring Automation & Division of Tasks

To look at the impact of automation, we focus on the rate of automation of steps belonging to the 1:1 case described in the previous section (so to keep the task content of each step constant across the human and machine scenarios). As multiple products in the HML dataset are used for this analysis, a superscript $p \in P$ is used to denote different products. We denote the motive power of step i (e.g. hand power, mule power, steam power, etc.) as ω_i^p. Next we construct an indicator of automation, i.e. a change in motive power between the hand and machine process. The HML dataset features no observations of motive power in hand processes such as steam or water shifting to "less mechanized"motive powers such as hand or animal power in the respective machine processes. Given this we treat all changes in motive power as a shift toward automation. Formally, for two 1:1 mapped steps $i^{H,p}$, $j^{M,p}$, the index of automation is given by the difference in motive power between the steps:

$$\theta = \begin{cases} 1 & \text{if} \quad \omega_i^p \neq \omega_j^p \\ 0 & \text{if} \quad \omega_i^p = \omega_j^p \end{cases} .$$

To look at the implications of $\bar{r}$ on the rate of automation, we construct a measure of the utilization of performers in each process step, $u = \frac{R}{r}$. The lower the utilization of performers, the lower the returns to increasing rate and the closer the performer is to $\bar{r}$. To compare between process steps which were or not automated, we use the parameters of a step's performer in the hand process to determine utilization given the volume in the machine process, as a proxy for $\bar{r}$. While the data does not include empirical production volume, we can recover an upper bound on the possible output of each process step: $R_j^{H,p} = r\mu$, where r is the rate of output per performer shift and μ is the number of performers demanded per shift. The maximum effective output of any step in a production process cannot be greater than the maximum output of every other step (bottlenecks), giving us $\overline{R}_i^{H,p} = \min_{i=1}^{i=N}(R_i^{H,p})$.

We next look at the division of tasks. To remove the effect of other changes beyond the division of tasks, we control for task content and for the level of automation. For the former we only consider the case of steps mapping from 1 to N (the decrease in the division of tasks is characterized by a M to 1 mapping. As fewer than 20 steps exhibit this property, we do not consider this scenario). To control for the level of automation we further restrict the sample selection to steps in which the motive power is unchanged.

C Additional Results

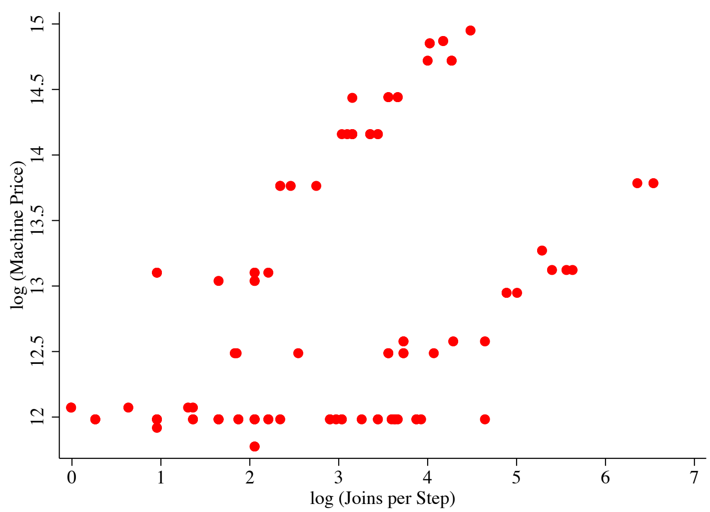


Figure C.1: Performer cost and step complexity. Prices in 2006 Dollars. For information on the data refer to Fuchs et al. (2008).