

The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality

David Hémous (University of Zurich and CEPR) and Morten Olsen (IESE)*

February 2016 (first draft: September 2013)

Abstract

This paper develops a framework for understanding the interplay between technological change and the income distribution by focusing on two types of innovations: the creation of new products, and the automation of existing tasks. We build an endogenous growth model in which the introduction of new products increases demand for labor across the skill distribution but automation reduces demand for low-skill labor at the product level. We solve for the full equilibrium path and show that the economy follows three phases: First, low-skill wages and therefore automation are low, while income inequality and the labor share are constant. Second, continuous horizontal innovation (the creation of new products) increases low-skill wages, which stimulates automation. This increases the skill-premium, depresses the growth rate of future low-skill wages (potentially to negative), and reduces the total labor share. Finally, the share of automated products stabilizes and low-skill wages grow at a positive but lower rate than high-skill wages. Surprisingly, a more productive automation technology benefits low-skill wages in the

*David Hémous, University of Zurich, david.hemous@econ.uzh.ch, Morten Olsen, IESE, molsen@iese.edu. Morten Olsen gratefully acknowledges the financial support of the European Commission under the Marie Curie Research Fellowship program (Grant Agreement PCIG11-GA-2012-321693) and the Spanish Ministry of Economy and Competitiveness (Project ref: ECO2012-38134). We thank Daron Acemoglu, Philippe Aghion, Ufuk Akcigit, Pol Antràs, Tobias Broer, Steve Cicala, Per Krusell, Brent Neiman, Jennifer Page, Andrei Shleifer, Che-Lin Su, Fabrizio Zilibotti and Joachim Voth amongst others for helpful comments and suggestions. We also thank seminar and conference participants at IIES, University of Copenhagen, Warwick, UCSD, UCLA Anderson, USC Marshall, Barcelona GSE Summer Forum, the 6th Joint Macro Workshop at Banque de France, Chicago Harris, the 2014 SED meeting, the NBER Summer Institute, the 2014 EEA meeting, Ecole Polytechnique, the University of Zurich, NUS, London School of Economics, the 2015 World Congress of the Econometric Society, ECARES, Columbia University, EIEF and Brown University. We thank Ria Ivandic and Marton Varga for excellent research assistance.

long-run by encouraging horizontal innovation. The model can account for some of the most salient features of modern labor markets, namely persistently increasing wage inequality, stagnating low-skill wages and wage polarization.

JEL: O41, O31, O33, E23, E25.

KEYWORDS: Endogenous growth, automation, horizontal innovation, directed technical change, income inequality.

1 Introduction

How does the automation of the production process drive economic growth and affect the distribution of income? Conversely, how do wages shape technological progress? The last 40 years in particular have seen dramatic changes in the income distribution with the skill premium rising throughout, low-skill wages stagnating and more recently a phase of wage polarization. These changes are often attributed to skill-biased technical change: by allowing for the use of machines in some tasks, automation increases economic output, but it also reduces the demand for certain types of labor, particularly low-skill labor. Autor, Levy and Murnane (2003) among others provide evidence in support of this mechanism. As the range of tasks that machines can perform has expanded considerably, the general public is increasingly worried about the negative consequences of technological progress. Yet, economists often argue that technological development also creates new products and tasks, which boost the demand for labor; and certainly many of today's jobs did not exist just a few decades ago.¹ Surprisingly, the economics literature lacks a dynamic framework to analyze the interaction between automation and the creation of new products. This paper provides the first model that can do so.

Of course, a large literature exists which relates exogenous technical change to the income distribution (e.g. Goldin and Katz, 2008, and Krusell, Ohanian, Ríos-Rull and Violante, 2002). Previous attempts at endogenizing the direction of technical change rely on factor-augmentation and exogenous shocks to the skill supply (Acemoglu, 1998). The novelty of our approach is that we present an endogenous growth version of a task framework in the vein of Autor, Levy and Murnane (2003) and Acemoglu and Autor (2011), in which the direction of innovation evolves endogenously.

The main lesson from our framework is the following. If tasks performed by a scarce factor (say labor) can potentially be automated but it is not presently profitable to do so, then, in a growing economy, the return to this factor will eventually increase sufficiently to make it profitable. Once automation has been triggered, the economy endogenously transitions from one aggregate production function to another and during this transition factor returns might drop. We characterize when this might happen and show that this decrease must be temporary. Although, we focus on a general equilibrium model with low-skill labor, the insights extend to subsectors of the economy and other scarce factors.

We consider an expanding variety growth model with low-skill and high-skill workers.

¹For instance, the introduction of the telephone led to the creation of new jobs. In 1970 there were 421 000 switchboard operators in the United States. This occupation has largely been automated today.

Horizontal innovation, modeled as in Romer (1990), increases the demand for both low- and high-skill workers. Automation allows for the replacement of low-skill workers with machines in production. It takes the form of a secondary innovation in existing product lines, similar to the secondary innovations in Aghion and Howitt (1996) (though their focus is on the interplay between applied and fundamental research, and not on automation). Within a firm, automation increases the demand for high-skill workers but reduces the demand for low-skill workers. “Non-automated” products only use low-skill and high-skill labor. Once invented, a specific machine is produced with the same technology as the consumption good.

We first take the level of technology as exogenous and show that an increase in the number of varieties will increase all wages. An increase in automation, however, will have the dual effect of increasing the overall productivity of the economy and allowing substitution away from low-skill workers, resulting in an ambiguous net effect on low-skill wages. Nevertheless, we show that for very general processes of horizontal and automation innovation, the asymptotic growth rate of low-skill wages must be positive, albeit strictly lower than that of high-skill wages.

Having studied the impact of technological change on wages, we require innovation to be the process of deliberate investment and show the key role played by low-skill wages. The cost advantage of an automated over a non-automated firm increases with the real wage of low-skill workers, so that the incentive to automate is not constant over time. As a consequence, the economy does not support a balanced growth path. Instead, an economy with an initially low level of technology first goes through a phase where growth is mostly generated by horizontal innovation and the skill premium and the labor share are constant. Only when low-skill wages are sufficiently high will firms invest in automation. During this second phase—where our model differs most from the existing literature—the share of automated products increases, the skill premium rises and, depending on parameters, the real low-skill wage may temporarily decrease. The total labor share decreases progressively, in line with recent evidence (Karabarbounis and Neiman, 2013). Finally, the economy moves towards its asymptotic steady-state. The share of automated products stabilizes as the entry of new, non-automated products compensates for the automation of existing ones. The total labor share stabilizes. Eventually, the economy will have endogenously shifted from a Cobb-Douglas aggregate production function to a nested CES one.

A simpler capital deepening model without automation innovation, but where low-

skill labor and capital are always substitutes would also feature a secular rise in the skill premium and a drop in the labor share. Yet our analysis shows that there are crucial differences between the two: First, contrary to a capital deepening model, we can generate stagnating or even declining real low-skill wages along the equilibrium path. Second, we can analyze the interaction between automation and another form of technological progress. For instance, we obtain the surprising result that a more productive automation technology increases the long-run growth rate of low-skill wages because it encourages horizontal innovation—although it may lead to lower low-skill wages for some time. This simple comparative statics result speaks to potential differences in policy implications between this setup and one without endogenous technical change. Third, by analyzing innovation patterns, our set-up makes some of the misconceptions that arise from restricting attention to factor-augmenting models apparent: for instance, intense automation can be consistent with a decline in labor productivity growth and the impact of automation on low-skill wages and the skill-premium need not be the strongest when expenses on automation are the largest; a response to two points of critique of the skill-biased technology change hypothesis put forward by Card and DiNardo (2002). In addition, while the elasticity of substitution between factors is of central importance for the labor share in factor-augmenting models (Piketty, 2014 and Karabarbounis and Neiman, 2014), our model highlights the role played by the share parameters.

Then, we extend the baseline model to include a supply response in the skill distribution, and calibrate it to match the evolution of the skill premium, the skill ratio, the labor share and productivity growth since the 1960s. As is common in the literature, for this exercise (and only this exercise) we identify skill groups with education groups, such that high-skill workers correspond to college-educated workers. This exercise demonstrates that our model is able to replicate the trends in the data quantitatively, even though we do not feed in any input time paths from the data as is usually done.

Finally, recent empirical work has increasingly found that workers in the middle of the income distribution are most adversely affected by technological progress. To address this, we extend the baseline model to include middle-skill workers as a separate skill-group. Products either rely on low-skill or middle-skill workers and the two skill-groups are symmetric except that automating to replace middle-skill workers is more costly (or alternatively machines are less productive in middle-skill firms). This implies that the automation of low-skill workers' tasks happens first, with a delayed automation process for the tasks of middle-skill workers. We show that this difference can reproduce

important trends in the United States income distribution: In a first period, there is a uniform dispersion of the income distribution, as low-skill workers' products are rapidly automated but middle-skill ones are not; while in the second period there is wage polarization: low-skill workers' share of automated products is stabilized, and middle-skill products are more rapidly automated.

Our modeling of automation as a skill-biased innovation is motivated by a large empirical literature. For instance, Autor, Katz and Krueger (1998) and Autor, Levy and Murnane (2003) use cross-sectional data to demonstrate that computerization is associated with relative shifts in demand favoring college-educated workers, Bartel, Ichniowski and Shaw (2007) present similar evidence at the firm level. Similarly, Graetz and Michaels (2015) show that the introduction of industrial robots leads to a reduction in the demand for low- (and middle-) skill workers. The idea that high wages could incentivize technological progress in the form of automation dates back to Habakkuk (1962). The empirical literature on this relationship is more modest, but Lewis (2001) finds that low-skill immigration slows down the adoption of automation technology and Hornbeck and Naidu (2014) find that the emigration of black workers from the American South favored the adoption of modern agricultural production techniques.

There is a small theoretical literature on labor-replacing technology. In Zeira (1998), exogenous increases in TFP raise wages and encourage the adoption of a capital-intensive technology analogous to automation in this paper. Acemoglu (2010) shows that labor scarcity induces innovation (the Habbakuk hypothesis), if and only if innovation is labor-saving, that is, if it reduces the marginal product of labor. Neither paper analyzes labor-replacing innovation in a fully dynamic model nor focuses on income inequality, as we do. Peretto and Seater (2013) build a dynamic model of factor-eliminating technical change where firms learn how to replace labor with capital. Since wages are constant the incentive to automate does not change over time. In addition, they do not focus on income inequality. Benzell, Kotlikoff, LaGarda and Sachs (2015), following Sachs and Kotlikoff (2012), build an overlapping generation model where a code-capital stock can substitute for labor. A technological shock which favors the accumulation of code-capital can lead to lower long-run GDP by reducing wages and thereby investment in physical capital. In both papers (and contrary to our model), the technological shock is completely exogenous. Finally, in work subsequent to our paper, Acemoglu and Restrepo (2015) also develop a growth model where technical changes involves automation and the creation of new tasks. We discuss their model in details in section 3.8.

A large literature has used skill-biased technical change (SBTC) as a possible explanation for the increase in the skill premium in developed countries since the 1970's, despite a large increase in the relative supply of skilled workers (see Hornstein, Krusell and Violante, 2005, for a more complete literature review). One can categorize theoretical papers into one of three strands. The first strand emphasizes the hypothesis of Nelson and Phelps (1966) that more skilled workers are better able to adapt to technological change (see Lloyd-Ellis, 1999; Caselli, 1999; Galor and Moav, 2000, and Aghion, Howitt and Violante, 2002). However, such theories mostly explain transitory increases in inequality whereas inequality has been increasing for decades. Our model, on the contrary, introduces a mechanism that creates permanent and widening inequality.

A second strand sees the complementarity between capital and skill as the source for the increase in the skill premium. Krusell, Ohanian, Ríos-Rull and Violante (2000) develop a framework where capital equipment and high-skill labor are complements. By adding the observed increase in the stock of capital equipment, they can account for most of the variation in the skill premium. Our model shares features with their framework: machines play an analogous role to capital equipment in their model, since they are more complementary with high-skill labor than with low-skill labor. The focus of our paper is different though since we seek to explain *why* innovation has been directed towards automation, and analyze the interactions between automation and horizontal innovation.

Finally, a third branch of the literature, building on Katz and Murphy (1992), considers technology to be either high-skill or low-skill labor augmenting and infers the bias of technology from changes in the relative supply and the skill-premium. Goldin and Katz (2008) employ this framework to conclude that technical change has been skill-biased throughout the 20th century in the United States (Katz and Margo, 2014, argue that the relative demand for white-collar workers has been increasing since 1820). This work, however, does not attempt to endogenize the skill bias of technical change. This is done in the (more theoretical) directed technical change literature (most notably Acemoglu, 1998, 2002 and 2007). Such models, which also use factor-augmenting technical change, deliver important insights about inequality and technical change, but they have no role for labor-replacing technology (a point emphasized in Acemoglu and Autor, 2011). In addition, even though income inequality varies, wages cannot decrease in absolute terms, and their asymptotic growth rates must be the same. The present model is also a directed technical change framework as economic incentives determine the form that technical change takes, but it deviates from the assumption of factor-augmenting

technologies and explicitly allows for labor-replacing automation, generating the possibility for (temporary) absolute losses for low-skill workers, and permanently increasing income inequality.

More recently, Autor, Katz, and Kearney (2006, 2008) and Autor and Dorn (2013), amongst others, show that whereas income inequality has continued to increase above the median, there has been a reversal below the median. They argue that the routine tasks performed by many middle-skill workers—storing, processing and retrieving information—are more easily done by computers than those performed by low-skill workers, now predominantly working in service occupations. This “wage polarization” has been accompanied by a “job polarization” as employment has followed the same pattern of decreasing employment in middle-skill occupations.² Our explanation is related but distinct: it is because low-skill tasks have already been heavily automated that automation is now more prominent in middle-skill tasks. Hence, we provide a unified explanation for the relative decline of middle-skill wages since the mid-1980s and the relative decline of low-skill wages in the period before.

The paper proceeds as follows: Section 2 describes the baseline model for exogenous technological change, it shows the consequences of technological change on wages and derives the asymptotic behavior. Section 3 endogenizes the path of technological change and describes the evolution of the economy through three phases. Section 4 calibrates an extended version of the model (with an endogenous labor supply response in the skill distribution) to the US economy since the 1960s. Section 5 extends the model to analyze wage polarization. Section 6 concludes.

2 A Baseline Model with Exogenous Innovation

2.1 Preferences and production

We consider a continuous time infinite-horizon economy populated by H high-skill and L low-skill workers. Both types of workers supply labor inelastically and have identical

²This phenomenon has also been observed and associated with the automation of routine tasks in Europe (Spitz-Oener, 2006, and Goos, Manning and Salomons, 2009). A related literature analyzes this non-monotonic pattern in inequality changes through the lens of assignment models where workers of different skill levels are matched to tasks of different skill productivities (e.g. Costinot and Vogel, 2010; Burstein, Morales and Vogel, 2014 and Feng and Graetz, 2015).

preferences over a single final good of:

$$U_{k,t} = \int_t^\infty e^{-\rho(\tau-t)} \frac{C_{k,\tau}^{1-\theta}}{1-\theta} d\tau,$$

where ρ is the discount rate, $\theta \geq 1$ is the inverse elasticity of intertemporal substitution and $C_{k,t}$ is consumption of the final good at time t by group $k \in \{H, L\}$. H and L are kept constant in our baseline model, but we consider the case where workers choose occupations based on relative wages and heterogeneous skill-endowments in Section 4.1.

The final good is produced by a competitive industry combining an endogenous set of intermediate inputs, $i \in [0, N_t]$ using a CES aggregator:

$$Y_t = \left(\int_0^{N_t} y_t(i)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}},$$

where $\sigma > 1$ is the elasticity of substitution between these inputs and $y_t(i)$ is the use of intermediate input i at time t . As in Romer (1990), an increase in N_t represents a source of technological progress. Throughout the paper, we use interchangeably the terms “intermediate input” and “product”.

We normalize the price of Y_t to 1 at all points in time and drop time subscripts when there is no ambiguity. The demand for each intermediate input i is:

$$y(i) = p(i)^{-\sigma} Y, \tag{1}$$

where $p(i)$ is the price of intermediate input i and the normalization implies that the ideal price index, $[\int_0^N p(i)^{1-\sigma} di]^{1/(1-\sigma)}$ equals 1.

Each intermediate input is produced by a monopolist who owns the perpetual rights of production. She can produce the intermediate input by combining low-skill labor, $l(i)$, high-skill labor, $h(i)$, and, possibly, type- i machines, $x(i)$, using the production function:

$$y(i) = \left[l(i)^{\frac{\epsilon-1}{\epsilon}} + \alpha(i) (\tilde{\varphi} x(i))^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon\beta}{\epsilon-1}} h(i)^{1-\beta}, \tag{2}$$

where $\alpha(i) \in \{0, 1\}$ is an indicator function for whether or not the monopolist has access to an automation technology which allows for the use of machines. If the product is not automated ($\alpha(i) = 0$), production takes place using a Cobb-Douglas production function with only low-skill and high-skill labor with a low-skill factor share of β . If the product

is automated ($\alpha(i) = 1$) machines can be used in the production process. We allow for perfect substitutability, in which case $\epsilon = \infty$ and the production function is $y(i) = [l(i) + \alpha(i)\tilde{\varphi}x(i)]^\beta h(i)^{1-\beta}$. The parameter $\tilde{\varphi}$ is the relative productivity advantage of machines over low-skill workers and G denotes the share of automated products.³

Since each input is produced by a single firm, from now on we identify each input with its firm and we refer to a firm which produces an automated product as an automated firm. We refer to the specific labor inputs provided by high-skill and low-skill workers in the production of different inputs as “different tasks” performed by these workers, so that each product comes with its own tasks. It is because $\alpha(i)$ is not fixed, but can change over time, that our model captures the notion that machines can replace low-skill labor in new tasks. A model in which $\alpha(i)$ were fixed for each product would only allow for machines to be used more intensively in production, but always for the same tasks.

Throughout the paper we will refer to x as “machines”, though our interpretation also includes any form of computer inputs, algorithms, the services of cloud-providers, etc. For simplicity, we consider that machines depreciate immediately, but Appendix 7.4 relaxes this assumption. Once invented, machines of type i are produced competitively one for one with the final good, so that the price of an *existing* machine for an automated firm is always equal to 1—hence technological progress in machine production follows that in the rest of the economy. Though a natural starting point, this is an important assumption and Appendix 7.3 presents a version of the model which relaxes it. Nevertheless, we stress that our model can capture the notion of a decline in the real cost of equipment: indeed automation for firm i can equivalently be interpreted as a decline of the price of machines i from infinity to 1.

2.2 Equilibrium wages

In this section we take the technological levels N (the number of products) and G (the share of automated products) as well as the employment of high-skill workers in production, H^P as given (we will let $H^P \leq H$ to accommodate later sections where high-skill labor is used to innovate). We now derive how wages are determined in equilibrium.

First, note that all automated firms are symmetric and therefore behave in the same

³Following existing evidence (Autor, Katz and Krueger, 1998, Autor, Levy and Murnane, 2003, or Bartel, Ichniowski and Shaw, 2007), we assume that the high-skill tasks cannot be automated. Yet, we discuss the automation of these tasks later in the paper.

way. Similarly all non-automated firms are symmetric. This gives aggregate output of:

$$Y = N^{\frac{1}{\sigma-1}} \times \left((1-G)^{\frac{1}{\sigma}} \underbrace{\left((L^{NA})^\beta (H^{P,NA})^{1-\beta} \right)}_{T_1}^{\frac{\sigma-1}{\sigma}} + G^{\frac{1}{\sigma}} \underbrace{\left(\left[(L^A)^{\frac{\epsilon-1}{\epsilon}} + (\tilde{\varphi}X)^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon\beta}{\epsilon-1}} (H^{P,A})^{1-\beta} \right)}_{T_2}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad (3)$$

where L^A (respectively L^{NA}) is the total mass of low-skill workers in automated (respectively non-automated) firms, $H^{P,A}$ (respectively $H^{P,NA}$) is the total mass of high-skill workers hired in production in automated (respectively non-automated) firms and $X = \int_0^N x(i)di$ is total use of machines. The aggregate production function takes the form of a nested CES between two sub-production functions. The first term T_1 captures the classic case where production takes place with constant shares between factors (low-skill and high-skill labor), while the second term T_2 represents the factors used within automated products and features the substitutability between low-skill labor and machines. G is the share parameter of the “automated” products nest and therefore an increase in G is T_2 -biased (as $\sigma > 1$). $N^{\frac{1}{\sigma-1}}$ is a TFP parameter. Besides the functional form the aggregate production function (3) differs from the often assumed aggregate CES production function in two ways: first, it is derived from the cost functions of individual firms and the technological change that we consider (new products and machines undertaking more tasks) is more concrete than the usual factor-augmenting technical change. Second, once we endogenize G we will be able to capture effects that the usual focus on an exogenous aggregate production function cannot.

The unit cost of intermediate input i is given by:

$$c(w, v, \alpha(i)) = \beta^{-\beta} (1-\beta)^{-(1-\beta)} \left(w^{1-\epsilon} + \varphi \alpha(i) \right)^{\frac{\beta}{1-\epsilon}} v^{1-\beta}, \quad (4)$$

where $\varphi \equiv \tilde{\varphi}^\epsilon$, w denotes low-skill wages and v high-skill wages. When $\epsilon < \infty$, $c(\cdot)$ is strictly increasing in both w and v and $c(w, v, 1) < c(w, v, 0)$ for all $w, v > 0$ (automation reduces costs). The monopolist charges a constant markup over costs such that the price is $p(i) = \sigma/(\sigma-1) \cdot c(w, v, \alpha(i))$.

Using Shepard’s lemma and equations (1) and (4) delivers the demand for low-skill

labor of a single firm.

$$l(w, v, \alpha(i)) = \beta \frac{w^{-\epsilon}}{w^{1-\epsilon} + \varphi \alpha(i)} \left(\frac{\sigma - 1}{\sigma} \right)^\sigma c(w, v, \alpha(i))^{1-\sigma} Y, \quad (5)$$

which is decreasing in w and v . The effect of automation on demand for low-skill labor in a given firm is generally ambiguous. This is due to the combination of a negative *substitution* effect (the ability of the firm to substitute machines for low-skill workers) and a positive *scale* effect (the ability of the firm to employ machines decreases overall costs, lowers prices and increases production). Since our focus is on labor-substituting innovation, we impose the condition $\epsilon > 1 + \beta(\sigma - 1)$ *throughout* the paper which is both necessary and sufficient for the substitution effect to dominate and ensure $l(w, v, 1) < l(w, v, 0)$ for all $w, v > 0$.

Let $x(w, v)$ denote the use of machines by an automated firm. The relative use of machines and low-skill labor for such a firm is then:

$$x(w, v)/l(w, v, 1) = \varphi w^\epsilon, \quad (6)$$

which is increasing in w as the real wage is also the price of low-skill labor relative to machines.

The iso-elastic demand (1), coupled with constant mark-up $\sigma/(\sigma - 1)$, implies that revenues are given by $R(w, v, \alpha(i)) = ((\sigma - 1)/\sigma)^{\sigma-1} c(w, v, \alpha(i))^{1-\sigma} Y$ and that a share $1/\sigma$ of revenues accrues to the monopolists as profits: $\pi(w, v, \alpha(i)) = R(w, v, \alpha(i))/\sigma$. Aggregate profits are then a constant share $1/\sigma$ of output Y , since output is equal to the aggregate revenues of intermediate inputs firms. We define $\mu \equiv \beta(\sigma - 1)/(\epsilon - 1) < 1$ (by our assumption on ϵ). Using (4), the relative revenues (and profits) of non-automated and automated firms are given by:

$$\frac{R(w, v, 0)}{R(w, v, 1)} = \frac{\pi(w, v, 0)}{\pi(w, v, 1)} = (1 + \varphi w^{\epsilon-1})^{-\mu}, \quad (7)$$

which is a decreasing function of w . Since non-automated firms rely more heavily on low-skill labor, their relative market share drops with higher low-skill wages.

The share of revenues in a firm accruing to high-skill labor in production is the same whether a firm is automated or not and given by $\nu_h = (1 - \beta)(\sigma - 1)/\sigma$. Using labor market clearing for high-skill workers in production ($\int_0^N h(i) di = H^P$) and aggregating

over those workers, we get that

$$vH^P = (1 - \beta) \frac{\sigma - 1}{\sigma} N [GR(w, v, 1) + (1 - G)R(w, v, 0)] = (1 - \beta) \frac{\sigma - 1}{\sigma} Y. \quad (8)$$

Using factor demand functions, the share of revenues accruing to low-skill labor is given by $\nu_l(w, v, \alpha(i)) = \frac{\sigma-1}{\sigma} \beta (1 + \varphi w^{\epsilon-1} \alpha(i))^{-1}$, and is lower for automated than non-automated firms. Similarly using low-skill labor market clearing ($\int_0^N l(i) di = L$), we obtain the aggregate revenues of low-skill workers as:

$$wL = N [GR(w, v, 1)\nu_l(w, v, 1) + (1 - G)R(w, v, 0)\nu_l(w, v, 0)]. \quad (9)$$

Taking the ratio of (8) over (9) and using (7) gives the following lemma.

Lemma 1. *For $\epsilon < \infty$, the high-skill wage premium is given by⁴*

$$\frac{v}{w} = \frac{1 - \beta}{\beta} \frac{L}{H^P} \frac{G + (1 - G)(1 + \varphi w^{\epsilon-1})^{-\mu}}{G(1 + \varphi w^{\epsilon-1})^{-1} + (1 - G)(1 + \varphi w^{\epsilon-1})^{-\mu}}. \quad (10)$$

For given L/H^P and $G > 0$, the skill premium is increasing in the absolute level of low-skill wages, which means that if G is bounded above 0, low-skill wages cannot grow at the same rate as high-skill wages in the long-run. This is the case because higher low-skill wages both induce more substitution towards machines in automated firms (as reflected by the term $(1 + \varphi w^{\epsilon-1})^{-1}$ in equation (10)) and improve the cost-advantage and therefore the market share of automated firms (term $(1 + \varphi w^{\epsilon-1})^{-\mu}$), which rely relatively less on low-skill workers.

With constant mark-ups, the cost equation (4) and the price normalization give:

$$\frac{\sigma}{\sigma - 1} \frac{N^{\frac{1}{1-\sigma}}}{\beta^\beta (1 - \beta)^{1-\beta}} (G(\varphi + w^{1-\epsilon})^\mu + (1 - G)w^{\beta(1-\sigma)})^{\frac{1}{1-\sigma}} v^{1-\beta} = 1. \quad (11)$$

This *productivity* condition shows the positive relationship between real wages and the level of technology given by N , the number of intermediate inputs, and G the share of automated firms. Together (10) and (11) determine real wages uniquely as a function of technology N, G and the mass of high-skill workers engaged in production H^P .

Though the production function implies that, at the firm level, the elasticity of substitution between high-skill labor and machines is equal to that between high-skill

⁴When machines and low-skill workers are perfect substitutes, $\epsilon = \infty$, the skill premium is given by $\frac{v}{w} = \frac{1-\beta}{\beta} \frac{L}{H^P}$ if $w < \tilde{\varphi}^{-1}$ such that no firm uses machines, and $\frac{v}{w} = \frac{1-\beta}{\beta} \frac{L}{H^P} \frac{G+(1-G)(\tilde{\varphi}w)^{-1}}{(1-G)(\tilde{\varphi}w)^{-1}}$ if $w > \tilde{\varphi}^{-1}$.

and low-skill labor, this does not imply that the same holds at the aggregate level. Therefore our paper is not in contradiction with Krusell et al. (2000), who argue that the aggregate elasticity of substitution between high-skill and low-skill labor is greater than the one between high-skill labor and machines.⁵

Given the amount of resources devoted to production (L, H^P) , the static equilibrium is closed by the final good market clearing condition:

$$Y = C + X \tag{12}$$

where $C = C_L + C_H$ is total consumption.

2.3 Technological change and wages

The consequences of technological changes on the level of wages are most easily seen with the help of Figure 1 which plots the skill-premium (10) and productivity (11) conditions in (w, v) space.⁶

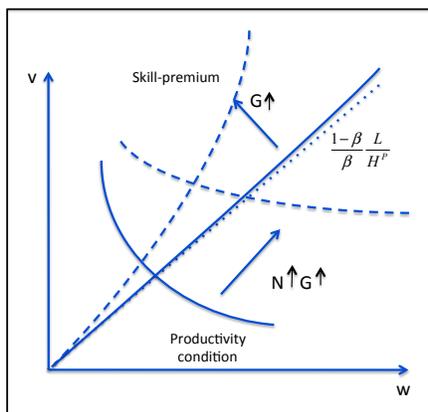


Figure 1: Evolution of high-skill (v) and low-skill (w) wages following technological changes. An increase in N pushes out the productivity condition increasing both wages. An increase in G pushes out the productivity condition and pivots the skill-premium counter-clockwise, which increases v but has an ambiguous effect on w .

An increase in the number of products, N , pushes out the productivity condition which results in an increase in both low-skill and high-skill wages. When $G = 0$, both

⁵This illustrates the difference between assuming elasticities of substitution of the aggregate production function and deriving them from elasticities of substitution at the firm level. In fact, in our model, the aggregate elasticities of substitution are not constant (see Appendix 8.3.5).

⁶When $\epsilon = \infty$ and $G = 1$, the productivity condition has an horizontal arm and the skill-premium condition a vertical one.

types of wages grow at the same rate as the skill premium condition is a straight line with slope $(1 - \beta)L/(\beta H^P)$. For $G > 0$, the skill premium curve is not linear and high-skill wages grow proportionally more than low-skill wages (following lemma 1).

An increase in the share of automated products G has a positive effect on high-skill wages and the skill premium with an ambiguous effect on low-skill wages. Indeed, higher automation increases the productive capability of the economy and pushes out the productivity condition (an *aggregate productivity* effect), but it also allows the economy to more easily substitute away from low-skill labor which pivots the skill-premium condition counter-clockwise (an *aggregate substitution*). The former effect increases low-skill wages and the latter decreases low-skill wages. Therefore automation is low-skill labor saving (in the vocabulary of Acemoglu, 2010) if and only if the aggregate substitution effect dominates the aggregate productivity effect. The following proposition (derived in Appendix 8.1) gives the full comparative statics.

Proposition 1. *Consider the equilibrium (w, v) determined by equations (10) and (11). It holds that*

A) An increase in the number of products N (keeping G and H^P constant) leads to an increase in both high-skill (v) and low-skill wages (w). Provided that $G > 0$, an increase in N also increases the skill premium v/w .

B) An increase in the share of automated products G (keeping N and H^P constant) increases the high-skill wages v and the skill premium v/w . Furthermore,

- *i) if $1/(1 - \beta) \leq \sigma - 1$, low-skill wages w are decreasing in G ;*
- *ii) if $\sigma - 1 < 1/(1 - \beta) < (\epsilon - 1)/\beta$, w is decreasing in G for N sufficiently low and inversely u-shaped in G otherwise, with $w|_{G=1} < w|_{G=0}$;*
- *iii) if $1/(1 - \beta) = (\epsilon - 1)/\beta$, w is inversely u-shaped in G with $w|_{G=1} = w|_{G=0}$;*
- *iv) if $(\epsilon - 1)/\beta < 1/(1 - \beta)$, w is increasing in G for N sufficiently low and inversely u-shaped otherwise (weakly if $\epsilon = \infty$), with $w|_{G=1} > w|_{G=0}$.*

The proposition states that $\beta/(1 - \beta) < \epsilon - 1$ — such that the elasticity of substitution between machines and low-skill workers is sufficiently large — is a necessary and sufficient condition to ensure that low-skill wages are lower in a world where all products are automated than in a world where none are. A low cost share of low-skill workers/machines, β , will make this more likely as automation then provides less cost savings and a lower aggregate productivity effect. For $\sigma - 1 < 1/(1 - \beta)$ and N (and therefore low-skill wages) sufficiently large, w is inversely u-shaped in G . This is because the automation of the first products has a relatively large productivity effect on

the economy, but a relatively small aggregate substitution effect since most firms are still non-automated, whereas the converse is true once G is sufficiently high.

In section 3, when we specify the innovation process, we will show that as the number of products N increases, the economy endogenously experiences a change in the share of automated products: from a low level, close to 0, to a higher level. As a result, growth will progressively become unbalanced with a rising skill premium, and for some parameter values, low-skill wages may temporarily decline.

2.4 Asymptotics for general technological processes

We study the asymptotic behavior of the model for given paths of technologies and mass of high-skill workers in production. For any variable a_t (such as N_t), we let $g_t^a \equiv \dot{a}_t/a_t$ denote its growth rate and $g_\infty^a = \lim_{t \rightarrow \infty} g_t^a$ if it exists. In Appendix 8.2.1 we derive

Proposition 2. *Consider three processes $[N_t]_{t=0}^\infty$, $[G_t]_{t=0}^\infty$ and $[H_t^P]_{t=0}^\infty$ where $(N_t, G_t, H_t^P) \in (0, \infty) \times [0, 1] \times (0, H]$ for all t . Assume that G_t , g_t^N and H_t^P all admit strictly positive limits. Then, the growth rates of high-skill wages and output admit limits with:*

$$g_\infty^v = g_\infty^Y = g_\infty^N / ((1 - \beta)(\sigma - 1)). \quad (13)$$

Part A) If $0 < G_\infty < 1$ then the asymptotic growth rate of w_t is given by

$$g_\infty^w = g_\infty^Y / (1 + \beta(\sigma - 1)). \quad (14)$$

Part B). If $G_\infty = 1$ and G_t converges sufficiently fast (more specifically if $\lim_{t \rightarrow \infty} (1 - G_t) N_t^{\psi(1-\mu)\frac{\epsilon-1}{\epsilon}}$ exists and is finite) then :

- If $\epsilon < \infty$ the asymptotic growth rate of w_t is positive at :

$$g_\infty^w = g_\infty^Y / \epsilon, \quad (15)$$

where $1 + \beta(\sigma - 1) < \epsilon$ by assumption.⁷

- If low-skill workers and machines are perfect substitutes then $\lim_{t \rightarrow \infty} w_t$ is finite and weakly greater than $\tilde{\varphi}^{-1}$ (equal to $\tilde{\varphi}^{-1}$ when $\lim_{t \rightarrow \infty} (1 - G_t) N_t^\psi = 0$).

This proposition first relates the growth rate of output and high-skill wages to the growth rate of the number of products. Without automation Y_t would be proportional

⁷If $\lim_{t \rightarrow \infty} (1 - G_t) N_t^{\psi(1-\mu)\frac{\epsilon-1}{\epsilon}} = \infty$ then $g_\infty^Y / \epsilon \leq g_\infty^w \leq g_\infty^Y / (1 + \beta(\sigma - 1))$.

to $N_t^{1/(\sigma-1)}$, as in a standard expanding-variety model: the higher the degree of substitutability between inputs the lower the gain in productivity from an increase in N_t . Here, because automation allows the use of machines as an additional input, there is an acceleration effect as the higher productivity also increases the supply of machines (as long as $G_\infty > 0$). Asymptotically, this effect is increasing in the factor share of low-skill workers/machines, β .⁸

Second, this proposition shows that, with positive growth in N_t , mild assumptions are sufficient to guarantee an asymptotic positive growth rate for w_t : in line with Proposition 1 the only case in which w_t would not be increasing occurs when G_t converges to 1 sufficiently fast and low-skill workers and machines are perfect substitutes. To gain further intuition, first consider the case in which $G_\infty < 1$. Since automated and non-automated products are imperfect substitutes, then so are machines and low-skill workers at the aggregate level. As the aggregate production function is a nested CES with asymptotically constant weights (see equation (3)), a growing stock of machines and a fixed supply of low-skill labor, implies that the relative price of a worker (w_t) to a machine (p_t^x) must grow at a positive rate. Since machines are produced with the same technology as the consumption good, $p_t^x = p_t^C$, where p_t^C is the price of the consumption good (1 with our normalization), and the real wage $w_t = w_t/p_t^C = (w_t/p_t^x)(p_t^x/p_t^C)$ must also grow at a positive rate.⁹

With growing wages, the relative market share of automated firms and their reliance on machines increase, which ensure that low-skill wages grow at a lower rate than the economy (with $G_\infty > 0$ the assumptions of Uzawa's theorem are not satisfied since horizontal innovation is not low-skill labor augmenting). Under our assumption that automation is labor-saving at the firm level ($\epsilon < 1 + \beta(\sigma - 1)$), the demand for low-skill labor increasingly comes from the non-automated firms. As a result, the ratio between high-skill and low-skill wages growth rates increases with a higher importance of low-skill workers (a higher β) or a higher substitutability between automated and non-automated products (a higher σ) since both imply a faster loss of competitiveness of the non-automated firms. On the other hand, it is independent of the elasticity of substitution between machines and low-skill workers, ϵ .

⁸Equation (13) also makes it clear why we must impose $\beta < 1$. If $\beta = 1$ the economy reaches a "world of plenty" in finite time, as infinite production is possible once the number of products N is sufficiently large relative to the productivity parameter $\tilde{\varphi}$. In reality one may think that other factors such as natural resources or land would then become the scarce factor.

⁹A generalized version of Proposition 2 is presented in Appendix 7.3 which allows for asymptotic (negative) growth in p_t^x/p_t^C and thereby potentially decreasing real wages for low-skill workers.

Now, consider the case of $G_\infty = 1$ and $\epsilon < \infty$ (and let convergence satisfy the condition in Part B of Proposition 2). Then an analogous argument demonstrates that low-skill wages must increase asymptotically, though the growth rate relative to that of the economy must be lower than when $G_\infty < 1$ as all firms are automated and automated firms more readily substitute workers for machines than the economy substitutes from non-automated to automated products. The more easily they substitute (the higher is ϵ) the lower the growth rate of low-skill workers wages. Only in the special case in which machines and low-skill workers are perfect substitutes in the production by automated firms and the share of automated firms is asymptotically 1 will there be economy-wide perfect substitution between low-skill workers and machines. In this case, w_t cannot grow asymptotically, but will still be bounded below by $\tilde{\varphi}^{-1}$, since a lower wage would imply that no firm would use machines.¹⁰

In general, the processes of N_t , G_t and H_t^P will depend on the rate at which new products are introduced, the extent to which they are initially automated, and the rate at which non-automated firms are automated. The following lemma derives condition under which $G_\infty < 1$, so that Part A of Proposition 2 applies.

Lemma 2. *Consider processes $[N_t]_{t=0}^\infty$, $[G_t]_{t=0}^\infty$ and $[H_t^P]_{t=0}^\infty$, such that g_t^N and H_t^P admit strictly positive limits. If i) the probability that a new product starts out non-automated is bounded below away from zero and ii) the intensity at which non-automated firms are automated is bounded above and below away from zero, then any limit of G_t must have $0 < G_\infty < 1$.*

Proof. See Appendix 8.2.2. □

In other words, as long as new non-automated products are continuously introduced (and stay non-automated for a non negligible time period), there will always be a share of non-automated products. These provide employment opportunities for low-skill workers which limits the relative losses of low-skill workers compared to high-skill workers (in that their wages grow according to (14) instead of (15)). This is endogenously what will happen in the full dynamic model that we now turn to.

¹⁰This provides one possible micro-foundation for the “Android Experiment” in Brynjolfsson and McAfee (2014) where an android is invented which can perform any task a human worker can do. Proposition 2 demonstrates that even if (asymptotically) this state is reached, low-skill workers will get the opportunity cost of such an android which in general will not tend to zero. Appendix 7.3 demonstrates that a necessary (though not sufficient) condition for low-skill wages to approach zero is that the cost of machines/androids falls faster than the consumption good.

Note that, the intuition given by the combination of Lemma 2 and Part A of Proposition 2 does not rely on new products being born identical to older products. In a model where new products are born more productive, the growth rate of high-skill wages and low-skill wages will obey equations (13) and (14), as long as the automation intensity is bounded and the economy grows at a positive but finite rate.

3 Endogenous innovation

We now model automation and horizontal innovation as a the result of intentional investment, which allows us to look at the impact of wages on technological change.

3.1 Modeling innovation

We assume that automation results from a risky investment: a non-automated firm which hires $h_t^A(i)$ high-skill workers in automation research, becomes automated according to a Poisson process with rate $\eta G_t^{\tilde{\kappa}} (N_t h_t^A(i))^\kappa$. Once a firm is automated it remains so forever. $\eta > 0$ denotes the productivity of the automation technology, $\kappa \in (0, 1)$ measures the concavity of the automation technology, $G_t^{\tilde{\kappa}}$, $\tilde{\kappa} \in [0, \kappa]$, represents possible knowledge spillovers from the share of automated products, and N_t represents knowledge spillovers from the total number of intermediate inputs. The spillovers in N_t are necessary to ensure that both automation and horizontal innovation can take place in the long-run.¹¹ Our set-up can be interpreted in two ways. From one standpoint, machines are intermediate input-specific and each producer needs to invent his own machine, which, once invented, is produced with the same technology as the consumption good.¹² From a sec-

¹¹The role of these spillovers is to compensate for the mechanical reduction in the amount of resources for automation that are available for each product when the number of product increases. With less spillovers (that is if the process depended on $N_t^q h_t^A$ with $q < 1$) automation would disappear as the amount of effective resources per firm available for automation ($N_t^q H/N_t$) would become arbitrarily small. With more spillovers ($q > 1$), the reverse occurs and firms could asymptotically get automated instantaneously. Furthermore, these spillovers can be micro-funded as follows: let there be a mass 1 of firms with N_t products (instead of assuming that each individual i is a distinct firm), then this functional form means that when a firm hires a mass $\tilde{H}_t^A(i)$ of high-skill workers in automation each of its non-automated products gets independently automated with a Poisson rate of $\eta G_t^{\tilde{\kappa}} (\tilde{H}_t^A(i)/(1 - G_t))^\kappa$.

¹²Alternatively, machine- i may be invented by an outside firm and then sold to the intermediate input producer. The rents from automation would then be divided between the intermediate input producer and the machine producer. Except for a constant representing the bargaining power of each party, it would not affect any of our results. Yet another alternative would be to have entrants undertaking automation and potentially displacing the original firm. This would not qualitatively affect the equilibrium as long as the incumbent has a positive probability of becoming automated.

ond standpoint, machines are produced by the final good sector, and each intermediate input producer must spend resources in adapting machines to his product line so as to make them substitutable with low-skill workers in a new set of tasks.

Horizontal innovation occurs in a standard manner. New intermediate inputs are developed by high-skill workers according to a linear technology with productivity γN_t , where $\gamma > 0$ is a productivity parameter.¹³ With H_t^D high-skill workers pursuing horizontal innovation, the mass of intermediate inputs evolves according to:¹⁴

$$\dot{N}_t = \gamma N_t H_t^D.$$

We assume that firms do not exist before their product is created. Coupled with our assumption that automation follows a continuous Poisson process, new products must then be born non-automated. This feature of the model is motivated by the idea that when a task is new and unfamiliar, the flexibility and outside experiences of workers allow them to solve unforeseen problems. As the task becomes routine and potentially codefiable a machine (or an algorithm) can perform it (as argued by Autor, 2013). In reality, some new tasks may be sufficiently close to older ones that no additional investment would be required to automate them immediately. Our results carry through if only a share of the new products are born non-automated and section 3.8 discusses an alternative set-up where automation is only undertaken at the entry stage.¹⁵

Therefore the rate and direction of innovation will depend on the equilibrium allocation of high-skill workers between production, automation and horizontal innovation. Defining the total mass of high-skill workers working in automation as $H_t^A \equiv \int_0^{N_t} h_t^A(i) di$, we get that high-skill labor market clearing leads to

$$H_t^A + H_t^D + H_t^P = H. \tag{16}$$

¹³As is common in the growth literature, in this set-up each firm is assumed to produce a good different from the others. Horizontal innovation, however, does not aim to represent the creation of new firms but the creation of new goods or services.

¹⁴The main results of this paper do not depend on the fact that firms are born non-automated. Appendix 8.6 presents a setting in which when a firm is born (and only when it is born), its owner can make it automated with probability $\min(\eta(N_t h_t^A)^\kappa, 1)$ by hiring h_t^A high-skill workers in automation. The transition of the economy is qualitatively identical. What is crucial is that higher low-skill wages increase the incentive to automate, not at what stage this automation takes place.

¹⁵The model predicts that the ratio of high-skill to low-skill labor *in production* is higher for automated than non-automated firms. However, this does not necessarily mean that automated firms have a higher ratio of high-skill to low-skill labor overall, since non-automated firms also hire high-skill workers for the purpose of automating. In particular, new firms do not always have a higher ratio of low to high-skill workers (and at the time of its birth a new firm only relies on high-skill workers).

3.2 Innovation allocation

We denote by V_t^A the value of an automated firm, by r_t the economy-wide interest rate and by $\pi_t^A \equiv \pi(w_t, v_t, 1)$ the profits at time t of an automated firm. The asset pricing equation for an automated firm is then given by

$$r_t V_t^A = \pi_t^A + \dot{V}_t^A. \quad (17)$$

This equation states that the required return on holding an automated firm, V_t^A , must equal the instantaneous profits plus appreciation. An automated firm only maximizes instantaneous profits and has no intertemporal investment decisions to make.

A non-automated firm has to decide how much to invest in automation. Denoting by V_t^N the value of a non-automated firm, we get the corresponding asset pricing equation:

$$r_t V_t^N = \pi_t^N + \eta G_t^{\tilde{\kappa}} (N_t h_t^A)^\kappa (V_t^A - V_t^N) - v_t h_t^A + \dot{V}_t^N, \quad (18)$$

where $\pi_t^N \equiv \pi(w_t, v_t, 0)$ and h_t^A is the mass of high-skill workers in automation research hired by a single non-automated firm (by symmetry $H_t^A = (1 - G_t) N_t h_t^A$). This equation has an analogous interpretation to equation (17), except that profits are augmented by the instantaneous expected gain from innovation $\eta G_t^{\tilde{\kappa}} (N_t h_t^A)^\kappa (V_t^A - V_t^N)$ net of expenditure on automation research, $v_t h_t^A$. This gives the first order condition:

$$\kappa \eta G_t^{\tilde{\kappa}} N_t^\kappa (h_t^A)^{\kappa-1} (V_t^A - V_t^N) = v_t, \quad (19)$$

which must hold at all points in time. The mass of high-skill workers hired in automation increases with the difference in value between automated and non-automated firms, and as such is increasing in current and future low-skill wages—all else equal.

Since non-automated firms get automated at Poisson rate $\eta G_t^{\tilde{\kappa}} (N_t h_t^A)^\kappa$, and since new firms are born non-automated, the share of automated firms obeys:

$$\dot{G}_t = \eta G_t^{\tilde{\kappa}} (N_t h_t^A)^\kappa (1 - G_t) - G_t g_t^N, \quad (20)$$

Free-entry in horizontal innovation guarantees that the value of creating a new firm cannot be greater than its opportunity cost:

$$\gamma N_t V_t^N \leq v_t, \quad (21)$$

with equality whenever there is strictly positive horizontal innovation ($\dot{N}_t > 0$).

The low-skill and high-skill representative households' problems are standard and lead to Euler equations which in combination give

$$\dot{C}_t/C_t = (r_t - \rho) / \theta, \quad (22)$$

with a transversality condition requiring that the present value of all time- t assets in the economy (the aggregate value of all firms) is asymptotically zero:

$$\lim_{t \rightarrow \infty} \left(\exp \left(- \int_0^t r_s ds \right) N_t \left((1 - G_t) V_t^N + G_t V_t^A \right) \right) = 0.$$

3.3 Equilibrium characterization

We define a feasible allocation and an equilibrium as follows:

Definition 1. A feasible allocation is defined by time paths of stock of products and share of those that are automated, $[N_t, G_t]_{t=0}^\infty$, time paths of use of low-skill labor, high-skill labor, and machines in the production of intermediate inputs $[l_t(i), h_t(i), x_t(i)]_{i \in [0, N_t], t=0}^\infty$, a time path of intermediate inputs production $[y_t(i)]_{i \in [0, N_t], t=0}^\infty$, time paths of high-skill workers engaged in automation $[h_t^A(i)]_{i \in [0, N_t], t=0}^\infty$, and in horizontal innovation $[H_t^D]_{t=0}^\infty$, time paths of final good production and consumption levels $[Y_t, C_t]_{t=0}^\infty$ such that factor markets clear ((16) holds) and good market clears ((12) holds).

Definition 2. An equilibrium is a feasible allocation, a time path of intermediate input prices $[p_t(i)]_{i \in [0, N_t], t=0}^\infty$, a time path for low-skill wages, high-skill wages, the interest rate and the value of non-automated and automated firms $[w_t, v_t, r_t, V_t^N, V_t^A]_{t=0}^\infty$ such that $[y_t(i)]_{i \in [0, N_t], t=0}^\infty$ maximizes final good producer profits, $[p_t(i), l_t(i), h_t(i), x_t(i)]_{i \in [0, N_t], t=0}^\infty$ maximize intermediate inputs producers' profits, $[h_t^A(i)]_{i \in [0, N_t], t=0}^\infty$ maximizes the value of non-automated firms, $[H_t^D]_{t=0}^\infty$ is determined by free entry, $[C_t]_{t=0}^\infty$ is consistent with consumer optimization and the transversality condition is satisfied.

We transform the system by introducing new variables for which the system of differential equations admits a steady-state. Specifically, we introduce $n_t \equiv N_t^{-\beta / [(1-\beta)(1+\beta(\sigma-1))]}$ and $\omega_t \equiv w_t^{\beta(1-\sigma)}$ which both tend towards 0 as N_t and w_t tend towards infinity. We define the normalized mass of high-skill workers in automation ($\hat{h}_t^A \equiv N_t h_t^A$), normalized high-skill wages and consumption ($\hat{v}_t = v_t N_t^{-\psi}$ and $\hat{c}_t = c_t N_t^{-\psi}$), where $\psi \equiv ((1-\beta)(\sigma-1))^{-1}$ (ψ is equal to the asymptotic elasticity of GDP with respect to N_t), and the variable $\chi_t \equiv \hat{c}_t^\theta / \hat{v}_t$ (χ_t is related to the mass of high-skill workers in pro-

duction). With positive entry in the creation of new products at all points in time, the equilibrium can then be characterized by a system of differential equations with two state variables n_t, G_t , two control variables, \hat{h}_t^A, χ_t and an auxiliary equation defining ω_t (see Appendix 7.1 for the derivation, in particular the system is given by equations (27), (28), (30) and (31)). We then get:

Proposition 3. *Assume that*

$$\rho \left(\frac{1}{\eta \kappa^\kappa (1 - \kappa)^{1-\kappa}} \left(\frac{\rho}{\gamma} \right)^{1-\kappa} + \frac{1}{\gamma} \right) < \psi H, \quad (23)$$

then the system of differential equations admits a steady-state $(n^, G^*, \hat{h}^{A*}, \chi^*)$ with $n^* = 0$, $0 < G^* < 1$ and positive growth $(g^N)^* > 0$.*

Proof. See Appendix 8.3.1. □

We will refer to the steady-state $(n^*, G^*, \hat{h}^{A*}, \chi^*)$ as an asymptotic steady-state for our original system of differential equations. In addition, the assumption that $\theta \geq 1$ ensures that the transversality condition always holds.¹⁶ For the rest of the paper we restrict attention to parameters such that there exists a unique saddle-path stable steady-state $(n^*, G^*, \hat{h}^{A*}, \chi^*)$ with $n^* = 0$, $G^* > 0$. Then, for an initial pair $(N_0, G_0) \in (0, \infty) \times [0, 1]$ sufficiently close to the asymptotic steady-state, the model features a unique equilibrium converging towards the asymptotic steady-state.¹⁷

In line with Lemma 2, the steady-state features $G^* \in (0, 1)$ as the automation intensity is positive but bounded. Therefore Proposition 2, part A applies: asymptotically high-skill wages grow faster than low-skill wages but the introduction of new non-automated products limits the ratio between the two growth rates.

¹⁶To see the intuition behind equation (23), consider the case in which the efficiency of the automation technology η is arbitrarily large, such that the model is arbitrarily close to a Romer model where all firms are automated. Then equation (23) becomes $\rho/\gamma < \psi H$, which mirrors the classical condition for positive growth in a Romer model with linear innovation technology. With a smaller η the present value of a new product is reduced such that the corresponding condition is more stringent

¹⁷Multiple asymptotic steady-states with $G^* > 0$ are technically possible but are not likely for reasonable parameter values (see Appendix 8.3.2). In addition, with two state variables (n_t and G_t) saddle path stability requires exactly two eigenvalues with positive real parts. In our numerical investigation, for all parameter combinations which satisfy the previous restrictions, this condition was always met.

3.4 Innovation incentives along the transitional path

A distinctive feature of this economy is that the path of technological change itself will be unbalanced through the transitional dynamics. In the following we think of the transitional path of the economy as going through three phases: a first phase where the incentive to automate is very low and the economy behaves close to a Romer model, a second phase in which automation pushes up G and the skill-premium condition of figure 1 pivots counter-clockwise and a final third phase where the economy approaches the steady state described by the previous section (a formal proof of the results in this section can be found in Appendix 8.3.3).

To see this, we combine (17) (18) and (19) to write the difference in value between an automated and a non-automated firm as:

$$r_t(V_t^A - V_t^N) = \pi_t^A - \pi_t^N - \frac{1 - \kappa}{\kappa} v_t h_t^A + (\dot{V}_t^A - \dot{V}_t^N).$$

Integrating over this equation (and using the transversality condition), we obtain that the difference in value between an automated and a non-automated firm is given by the discounted flow of the difference in profits adjusted for the cost of automation and the probability of getting automated:

$$V_t^A - V_t^N = \int_t^\infty \exp\left(-\int_t^\tau r_u du\right) \left(\pi_\tau^A - \pi_\tau^N - \frac{1 - \kappa}{\kappa} v_\tau h_\tau^A\right) d\tau. \quad (24)$$

The rate of automation depends on the normalized mass of high-skill workers in automation ($\hat{h}_t^A = N_t h_t^A$), which following (19) depends on the ratio between the gain in firm value from automation and the high-skill wage divided by the number of products:

$$\hat{h}_t^A = \left(\kappa \eta G_t^{\tilde{\kappa}} \frac{V_t^A - V_t^N}{v_t / N_t}\right)^{1/(1-\kappa)}. \quad (25)$$

Crucially as the number of products in the economy increases, the right-hand side of this expression will change value drastically.

First Phase. High-skill wages v_t and aggregate profits are of the same order (both are proportional to output). When the number of products N_t is low, wages, including low-skill wages are low and therefore automated and non-automated firms have similar profits. As a result aggregate profits are close to $N_t \pi_t^N$ so that π_t^N and v_t / N_t are of the same order (and grow at similar rates). Recalling that (7) gives

$\pi_t^A - \pi_t^N = (1 + \varphi w_t^{\varepsilon-1})^\mu \pi_t^N$, we get that as long as w_t is small relative to $\tilde{\varphi}^{-1}$, then $(\pi_t^A - \pi_t^N) / (v_t/N_t)$ must be small too.¹⁸ With a positive discount rate, this implies that $(V_t^A - V_t^N) / (v_t/N_t)$ and therefore \hat{h}_t^A must also be small (see Appendix 8.3.3 for a proof). Hence the economy initially experiences little automation.

Therefore, for sufficiently low initial value of N_t the behavior of the economy is close to that of a Romer model with a Cobb-Douglas production function between low-skill and high-skill labor. Economic growth is driven by horizontal innovation and the skill premium and the factor shares are nearly constant. Naturally, if G_t is not initially low, it must depreciate during this period following equation (20). This corresponds to what we label as the first phase of the economy.

Second Phase. As horizontal innovation continuously increases low-skill wages (for low N_t , w_t grows at the rate $g_t^N / (\sigma - 1)$) the approximation derived on the basis that w_t is small relative to $\tilde{\varphi}^{-1}$ becomes progressively worse. The gain from automating from $(V_t^A - V_t^N) / (v_t/N_t)$ increases and ceases to be a low-number.¹⁹ Without the externality in the automation technology ($\tilde{\kappa} = 0$), it is then direct from (25) that the mass of high-skill workers devoted to automation becomes significantly different from 0, so that the Poisson rate of automation $\eta (N_t h_t^A)^\kappa$ increases and so does the share of automated products G_t . For $\tilde{\kappa} > 0$, the depreciation in the share of automated products during the first phase might gradually makes the automation technology less effective which can delay or even potentially prevent the take-off of automation.²⁰

In line with Proposition 1, both the increase in G_t but also the increase in the number of products (now with G_t substantially higher than 0) lead to an increase in the skill premium. We will label this time period where the share of automated products in the economy increases sharply the second phase (the transition between phases is smooth and therefore the exact limits are arbitrary). Arguably, this time period is the one where our model differs the most from the rest of the literature.

Third Phase. With a share of automated products significantly different from 0, aggregate profits and the profits of an automated firm multiplied by the number of products ($N_t \pi_t^A$) must be of the same order. Recalling that high-skill wages and aggregate

¹⁸More specifically one finds that $\frac{\pi_t^A - \pi_t^N}{v_t/N_t} = \psi H_t^P \frac{((1 + \varphi w_t^{\varepsilon-1})^\mu - 1)}{1 + G_t((1 + \varphi w_t^{\varepsilon-1})^\mu - 1)}$, which is small if $\varphi w_t^{\varepsilon-1}$ is small and increasing in w_t (and therefore in N_t) for given G_t and H_t^P .

¹⁹When G_t is low, we still have that $N_t \pi_t^N$ and v_t are of the same order and as a result the ratio $(V_t^A - V_t^N) / (v_t/N_t)$ grows like $(\pi_t^A - \pi_t^N) / \pi_t^N = (1 + \varphi w_t^{\varepsilon-1})^\mu - 1$.

²⁰Automation does still take off if either G_0 and N_0 are not too low or, for any values of $N_0, G_0 > 0$ whenever $1 - \kappa - \tilde{\kappa} > 0$ —see Appendix 8.3.4. Finally, if we were to assume instead that the automation technology is given by $\min \{ \eta G_t^{\tilde{\kappa}}, \eta \} (N_t h_t^A)^\kappa$, then automation would always take off.

profits are both proportional to output, this ensures that the ratio $(V_t^A - V_t^N) / (v_t/N_t)$ remains bounded. As a result the normalized mass of high-skill workers in automation research $(N_t h_t^A)$ also stays bounded (see (25)). In line with Lemma 2 a bounded Poisson rate of automation ensures that the share of automated products stabilizes below 1. Therefore, the economy will experience a third phase where the share of automated products is approximately constant. Following Proposition 1, both high-skill and low-skill wages grow but high-skill wages grow at a higher rate. Though in Phase 3 the share parameters of the nested CES function are constant, the model continues to differ from a generic capital deepening model in that long-run growth is endogenized and depends on its interaction with automation (in particular, see Proposition 4 below).

3.5 An illustration of the transitional dynamics

In order to illustrate the previous results and to further analyze the behavior of our economy under various parametric assumptions, we now turn to numerical simulations.²¹ The following section will then relate our theoretical results to the historical experience of the US economy. Table 1 presents our baseline parameters. Section 4 employs Bayesian techniques to estimate the parameters, but the focus of this section is theoretical and we simply choose “reasonable” parameters (see Appendix 7.2.4 for a systematic exploration of the parameter space). As our goal is to characterize the evolution of an economy which transitions from automation playing a small to a central role, we choose an initially low level of automation ($G_0 = 0.001$) and an initial mass of intermediate inputs small enough to ensure that the real wage is initially low relative to the productivity of machines. This ensures that the economy will start in the ‘first phase’ described above (with a higher (N_0, G_0) the economy may directly start in the second or third phases).

Table 1: Baseline Parameter Specification

σ	ϵ	β	H	L	θ	η	κ	$\tilde{\varphi}$	ρ	$\tilde{\kappa}$	γ
3	4	2/3	1/3	2/3	2	0.2	0.5	0.25	0.02	0	0.3

Baseline Parameters. The time unit is 1 year. Total stock of labor is 1 and we set $L = 2/3$ and $\beta = 2/3$ such that absent automation and if all high-skill workers were in production the skill premium would be 1. The initial mass of products is $N_0 = 1$ and the productivity parameter for machines is $\tilde{\varphi} = 0.25$, which ensures that at $t = 0$, the

²¹We employ the so-called “relaxation” algorithm for solving systems of discretized differential equations (Trimborn, Koch and Steger, 2008). See Appendix 8.4 for details.

cost advantage of automated firms is very small (their profits are 0.004% higher). We set $\sigma = 3$ to capture an initial labor share close to $2/3$. The elasticity of substitution between machines and low-skill workers in automated firms is $\epsilon = 4$. The innovation parameters (γ, η, κ) are chosen such that GDP growth is close to 2% both initially and asymptotically, and we first consider the case where there is no externality from the share of automated products in the automation technology, $\tilde{\kappa} = 0$ —hereafter, we will refer to this externality as the externality in automation technology (although there is also an externality from the total mass of products). The parameters ρ and θ are chosen such that the interest rate is around 6% (at the beginning and at the end of the transition).

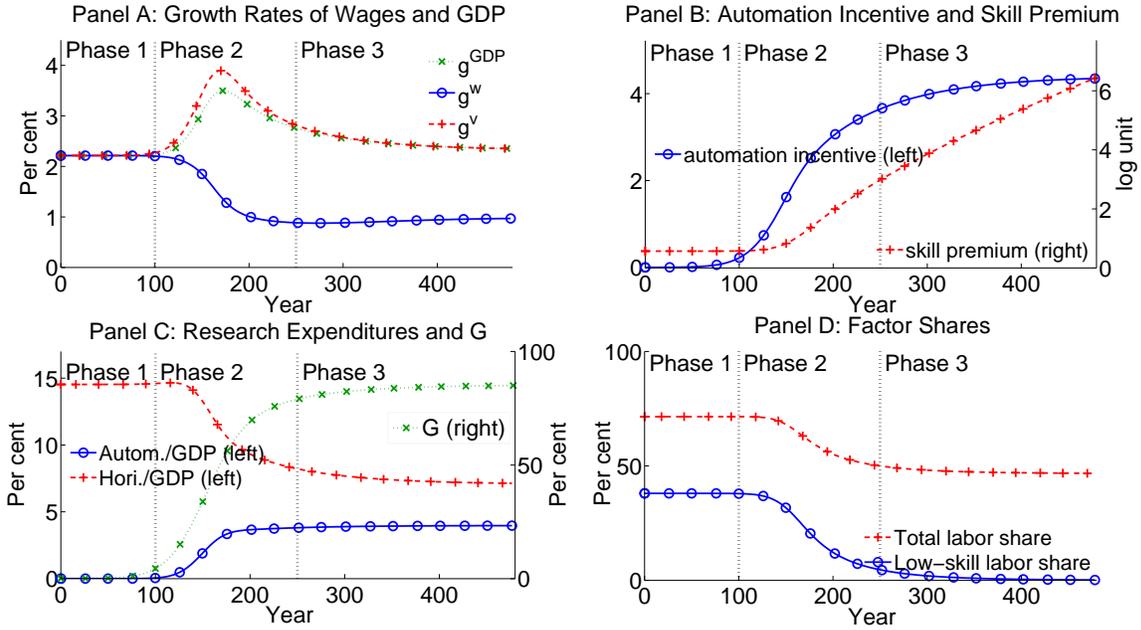


Figure 2: Transitional Dynamics for baseline parameters. Panel A shows growth rates for GDP, low-skill wages (w) and high-skill wages (v), Panel B the incentive to automate, $(V_t^A - V_t^N) / (v_t/N_t)$, and the skill premium, Panel C the total spending on horizontal innovation and automation as well as the share of automated products (G), and Panel D the wage share of GDP for total wages and low-skill wages.

Figure 2 plots the evolution of the economy. Based on the behavior of G_t (Panel C) we delimit Phase 1 as corresponding to the first 100 years and Phase 2 as the period between year 100 and year 250.

Innovation. As previously derived and as shown in Panel C firms initially spend very little on automation and the share of automated firms, G_t , stays initially very close to 0. This occurs because with a low initial level of N_t , low-skill wages are low, so that the

gains from automation, $V_t^A - V_t^N$, are very low relative to the high-skill wage normalized by the number of products (v_t/N_t)—as shown in Panel B. With growing low-skill wages, the incentive to automate picks up a bit before year 100. Then, the economy enters the second phase as automation expenses sharply increase (up to 4% of GDP), leading to an increase in the share of automated products G_t . Despite a high level of expenditures in automation, the share of automated products eventually stabilizes in the third phase below 1, as the constant arrival of new, non-automated products depreciates it.

As shown in Panel C, spending on horizontal innovation as a share of GDP declines during Phase 2 and ends up being lower in Phase 3 than Phase 1. Intuitively, horizontal innovation becomes less interesting because new (non-automated) products will have to compete with increasingly productive automated firms and therefore get a smaller initial market share. Though this is not a general feature of our model, such possibility distinguishes our paper from the ‘Habbakuk hypothesis’ literature: high level of wages may encourage some innovation (automation), but it may also discourage other forms of innovation, leading to an ambiguous overall impact on growth.

Panel A shows that *GDP* growth is the highest in the middle of Phase 2 and roughly the same in Phases 1 and 3: the rate of horizontal innovation is lower in Phase 3 but this is compensated by the elasticity of *GDP* wrt. N_t being higher (at $1/(\sigma - 1)$ instead of $1/[(\sigma - 1)(1 - \beta)]$). As a result, the phase of intense automation (when the share of automated products increases) is associated with a temporary boost of growth.

Wages. In the first phase, and referring to figure 1, the skill-premium conditions stays close to the straight dotted line with slope $\frac{1-\beta}{\beta} \frac{L}{HP}$ associated with a Cobb-Douglas production. Continuous horizontal innovation pushes the productivity condition towards the North East so that both wages grow at around 2% (Panel A).

As rising low-skill wages trigger the second phase, the skill-premium condition pivots counter-clockwise and bends upwards increasing the growth rate of high-skill wages to almost 4% and suppressing the growth rate of low-skill wages to around 1%. (since there are no financial constraints, the two types share a common consumption growth rate throughout, see Appendix 7.2.1). For our specific parameter choice (satisfying the conditions of Proposition 1 B.ii)), increases in G_t have a negative impact on w_t throughout the transition, but it is sufficiently slow relative to the increase in N_t that low-skill wages grow at a positive rate throughout—which section 3.7 demonstrates need not be the case. It is precisely this movement of the skill-premium curve that an alternative model with constant G (i.e. one where the fraction of tasks that can be performed

with machines is constant) or a capital deepening model could not reproduce, and consequently such a model would not feature labor-saving innovation. In addition to the effects of changing G_t and N_t , changes in the mass of high-skill workers in production, H_t^P , affect the skill premium, but these effects are quantitatively dominated. In the third phase, the skill-premium condition no longer moves, but the continuous rise in the productivity condition continues to increase the skill-premium.

Capital and labor shares. Panel D of figure 2 plots the labor share and the low-skill labor share. To understand their evolution, first note that $GDP = Y - X + v(H^D + H^A)$, as GDP includes R&D investments done by high-skill labor, but not intermediate inputs, X . Since machines are not part of the capital stock in this baseline version of the model (see Appendix 7.4 for an alternative specification), capital income corresponds to aggregate profits. Due to a constant markup, these profits are a constant share of output. During the first phase, X is low and $v(H^D + H^A)$ are roughly a constant share of output, implying a nearly constant capital share. With low-skill and high-skill wages growing at the same rate, the low-skill share is also constant during this period.

With the advent of automation in Phase 2, the increased use of intermediate inputs implies a decreasing GDP/Y and profits and thereby the capital share becomes a growing share of GDP . Working contrary to this is the increase in innovation as only high-skill workers work in innovation, but the net effect is an increase in the capital share. For different parameter values, the drop in the labor share can be delayed relative to the rise in the skill premium (see Appendix 7.2.2). As the growth rate of low-skill wages starts declining and falls to a level permanently lower than that of GDP, the low-skill labor share declines eventually to approach zero. The high-skill wage share, however, asymptotically grows at the rate of GDP and stabilizes at a higher value than in Phases 1 and 2. The ratio of wealth to GDP also increases during Phase 2 and asymptotes a constant in Phase 3 (see Appendix 7.2.1).

Elasticity of substitution. At the aggregate level, our model boils down to a nested CES production function (see equation (3)), and Phase 2 corresponds to a period where the share parameter of the composite which features substitutability between machines and low-skill labor, G_t , rises. This change in the share parameter is micro-founded and receives a very natural interpretation, which is precisely the advantage of a task framework. In contrast, the academic debate on income distribution often focuses on the value of the elasticity of substitution between different factors. Here the value of the aggregate elasticity of substitution does not play the central role. In fact, the

Morishima’s elasticity of substitution between low-skill labor and machines (or machines and low-skill labor) actually declines in Phase 2 from a value close to ϵ to a value close to $1 + \beta(\sigma - 1)$ (see Appendix 8.3.5).

Growth decomposition. Figure 3 performs a growth decomposition exercise for low-skill and high-skill wages by separately computing the instantaneous contribution of each type of innovation. We do so by performing the following thought experiment: at a given instant t , for given allocation of factors, suppose that all innovations of a given type fail. By how much would the growth rates of w and v change? This exercise is complementary to the one performed in Figure 1 which focuses on the impact of technological levels instead of innovations.²² In Phase 1, there is little automation, so wage growth for both skill-groups is driven almost entirely by horizontal innovation. In Phase 2, automation sets in. Low-skill labor is then continuously reallocated from existing products which get automated, to new, not yet automated, products. The immediate impact of automation on low-skill wages is negative, while horizontal innovation has a positive impact, as it both increases the range of available products and decreases the share of automated products. The figure also shows that automation plays an increasing role in explaining the growth rate of high-skill wages, while the contribution of horizontal innovation declines. This is because new products capture a smaller and smaller share of the market and therefore do not contribute much to the demand for high-skill labor. Consequently, automation is skill-biased while horizontal innovation is unskilled-biased.²³ We stress that this growth decomposition is for changes in the rate of automation and horizontal innovation at a given point in time. This should not be interpreted as “automation being harmful” to low-skill workers in general. In fact, as we demonstrate in Section 3.7, an increase in the effectiveness of the automation technology, η , will have positive long-term consequences. A decomposition of g_t^{GDP} would look

²²More specifically we can write $w_t = f(N_t, G_t, H_t^P)$, using equations (10) and (11). Differentiating with respect to time and using equation (28) gives:

$$g_t^w = \left(\frac{N_t}{w_t} \frac{\partial f}{\partial N} - \frac{G_t}{w_t} \frac{\partial f}{\partial G} \right) \gamma H_t^D + \frac{1}{w_t} \frac{\partial f}{\partial G} \eta G_t^{\bar{\kappa}} (1 - G_t) (\hat{h}_t^A)^{\kappa} + \frac{1}{w_t} \frac{\partial f}{\partial H^P} \dot{H}_t^P.$$

Figure 3 plots the first two terms as the growth impact of expenses in horizontal innovation and automation, respectively. The third term ends up being negligible. We perform a similar decomposition for v_t .

²³Proposition 1 shows analytically that automation is skilled-biased and that an increase in N at a given $G > 0$ is also skilled-biased. Horizontal innovation corresponds to an increase in the number of non-automated products, that is an increase in N but keeping GN constant. One can show analytically that horizontal innovation is unskilled-biased when w is high enough (which is obtained for $\epsilon < \infty$ and N large enough), but might be skill biased for low N .

similar to the decomposition of g_t^v : while growth is initially almost entirely driven by horizontal innovation, automation becomes increasingly important in explaining growth.²⁴

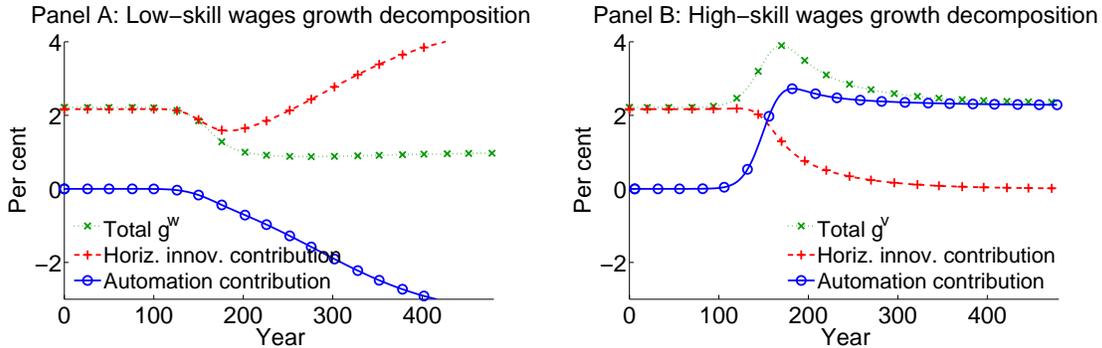


Figure 3: Growth decomposition. Panel A: The growth rate of low-skill wages and the instantaneous contribution from horizontal innovation and automation, respectively. Panel B is analogous for high-skill wages. See text for details.

3.6 Comparison with the historical experience

We now relate the qualitative lessons of our model to historical experience. Although, when undertaking our quantitative analysis in section 4, we will focus on the last 40 years, we believe that the forces at work in our model were relevant much before.

Secular increase in the relative skill demand. Our model predicts a continuous increase in the skill premium from Phase 2. It is a well established fact that the college premium (considered to be a good proxy for the skill premium over that time period) has been steadily increasing in the United States since the 1980s. Moving back in time, the skill premium experienced periods of decline (such as the 1970s) but these can be accounted for by exogenous changes in the relative supply of skills which our models does not capture. Factoring in the evolution on the supply side, Goldin and Katz (2008) show that technological change has been skill-biased throughout the 20th century. Even before, Katz and Margo (2014) argue that the relative demand for highly skilled workers (in professional, technical and managerial occupations) has increased steadily from perhaps as early as 1820 to the present.

This contrasts our paper with most of the growth literature which features a balanced growth path and therefore does not have permanently increasing labor inequality. For instance, in Acemoglu (1998), low-skill and high-skill workers are imperfect substitutes in

²⁴This is about instantaneous growth, as shown in Proposition 2, long-run growth is ultimately determined by the endogenous rate of horizontal innovation.

production. Yet, since the low-skill augmenting technology and the high-skill augmenting technology grow at the same rate asymptotically, the relative stocks of effective units of low-skill and high-skill labor is constant, leading to a constant relative wage.

Furthermore, there is no simple one-to-one link between automation spending and rising inequality in our model. Here, automation spending is higher in Phase 3 than in Phase 2 (Panel C in figure 2), yet the growth in the skill premium is slower. Card and DiNardo (2002) argue that inequality rising the most in the early to mid 1980s, and technological change continuing since squares poorly with the predictions of a framework based on skill-biased technological change. This, in fact, is in line with our model.²⁵

Of course, the definition of who is ‘high’ skill and who is ‘low’ skill matters. Hence, the mechanization of the 19th century, which replaced skilled artisans, or the computerization of the last 30 years, did not aim at replacing the most unskilled workers. Therefore our model provides a good account of the historical experience only if the definition of ‘high-skill’ is restrictive. The next section, which introduces a group of middle-skill workers, helps us account for such events.

Capital and labor shares. Our model also predicts a slow drop in the labor share during most of Phase 2 (and a rise in the capital to income ratio). This is consistent with Karabarounis and Neiman (2013) who find a global reduction of 5 percentage points in labor’s share of corporate gross value added over the past 35 years, and with Elsby, Hobijn and Sahin (2013) who find similar results for the United States.

However, the capital share of income and the wealth to income ratio have followed a U-curve in the 20th century (Piketty and Zucman, 2014 and Piketty, 2014). Although a small temporary decline in the capital share at the beginning of Phase 2 can be accounted for by the model (see Appendix 7.2.2), such large movements cannot. The early decline in the capital share is at least partly due to the two World Wars, changes in the tax system and the structural shift away from the agricultural sector, which this model does not capture. The latter, could be captured with a nested structure with an elasticity of substitution between broad sectors of less than 1. If these sectors differ in how easy they are to automate, then intense automation will happen sequentially. As this happens spending shares in non-automated sectors would increase (as in Acemoglu and Guerrieri, 2008) securing a higher growth rate for low-skill wages. This would replicate the broad features of an economy switching from agriculture, to manufacturing, and then services.²⁶

²⁵Intuitively the elasticity of the skill-premium with respect to the skill-bias of technology is not constant in our model, contrary to a CES framework with factor-augmenting technologies.

²⁶In addition, at the product level, an elasticity of substitution lower than 1 between the low-

3.7 Interaction between horizontal innovation and automation

We now investigate the interaction between horizontal innovation and automation by changing the innovation parameters. This allows us to further distinguish our model with endogenous technological change from alternative models. Appendix 7.2.4, carries out a more systematic comparative statics exercise.

Declining low-skill wages. Empirical evidence suggests that low-skill wages have been stagnating and perhaps even declining in recent periods. Because our model features a labor saving innovation, it can accommodate declining low-skill wages in Phase 2. This could not happen if the share of automated products were fixed, or in a capital deepening model with perfectly elastic capital, and it is not the case in the previous DTC literature either. In this model it happens when the skill-premium pivots sufficiently fast counter-clockwise compared with the movement of the productivity condition in figure 1. We ensure this in figure 4 by setting $\tilde{\kappa} = 0.49$, thereby introducing the externality in automation.²⁷ Initially G_t is small and the automation technology is quite unproductive. Hence, Phase 2 starts later, even though the ratio $(V_t^A - V_t^N) / (v_t/N_t)$ has already significantly risen (Panel B). Yet, Phase 2 is more intense once it gets started, partly because of the sharp increase in the productivity of the automation technology (following the increase in G_t) and partly because low-skill wages are higher. Intense automation puts downward pressure on low-skill wages. At the same time, horizontal innovation drops considerably, both because new firms are less competitive than their automated counterparts, and because the high demand for high-skill workers in automation innovations increases the cost of inventing a new product. This results in a short-lived decline in low-skill wages. Indeed, the decline in w_t (and increase in high-skill wage v_t) lowers the incentive to automate (Panel B), which in return reduces automation. This reflects two general points: First, just as increases in w_t encourage automation; reductions in w_t discourage it. Second, decreases in wages require two conditions to be met: First, it must be technically possible for low-skill wages to decrease, which here can be done by increases in G_t . Second, it must be privately optimal for agents to choose such a path for G_t . Here this condition can be met by either of two assumptions: one, the externality in automation ($\tilde{\kappa} > 0$) and second the fact that innovation automation has to be paid as an upfront cost instead of a fixed cost every period (and consequently low-skill wages

skill/machines aggregate and high-skill workers would secure a higher level for the labor share.

²⁷We choose this value for $\tilde{\kappa}$ instead of 0.5, because in that case there is no horizontal innovation for some time periods (that is (21) holds with a strict inequality). This is not an issue in principle but simulating this case would require a different numerical approach.

can drop for $\tilde{\kappa} = 0$ — albeit for a small parameter set — see Appendix 7.2.3).

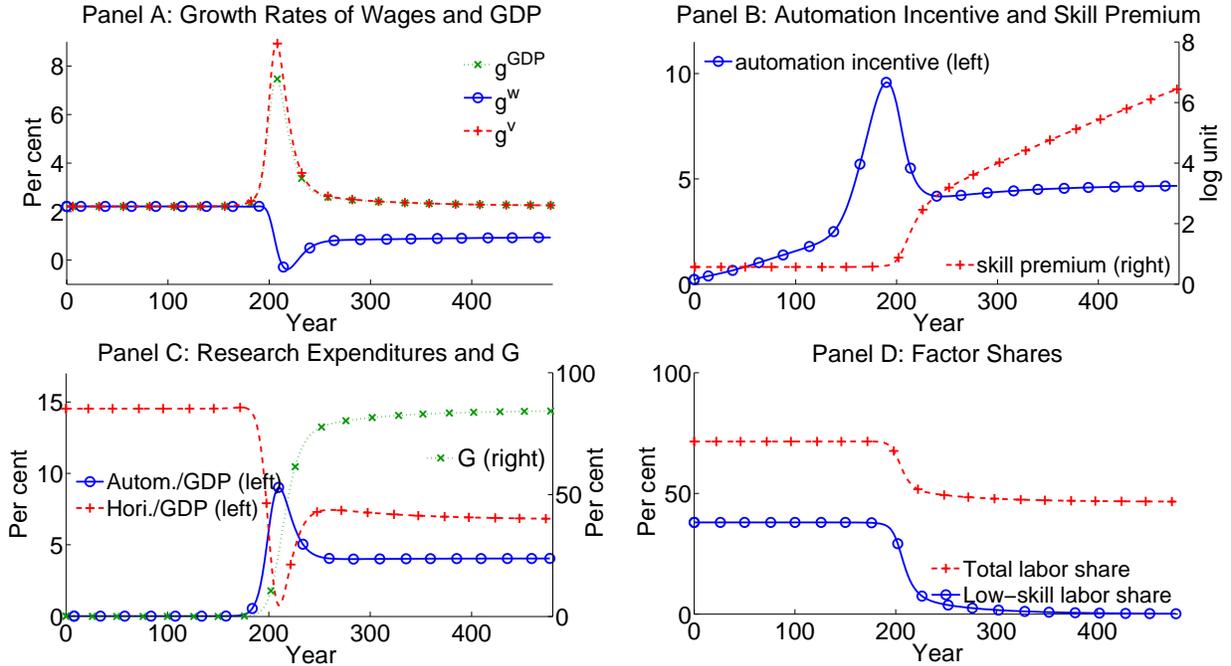


Figure 4: Transitional Dynamics. Note: same as for Figure 2 but with an automation externality of $\tilde{\kappa} = 0.49$.

Innovation parameters. Figure 5 shows the impact (relative to the baseline case) of increasing productivity in the automation technology to $\eta = 0.4$ (from 0.2) and the productivity in the horizontal innovation technology to $\gamma = 0.32$ (instead of 0.3). A higher η initially has no impact during Phase 1, but it moves Phase 2 forward as investing in automation technology is profitable for lower level of low-skill wages. Since automation occurs sooner, the absolute level of low-skill wages drops relative to the baseline case (Panel B), which leads to a fast increase in the skill premium. However, as a higher η means that new firms automate faster, it encourages further horizontal innovation. A faster rate of horizontal innovation implies that the skill premium keeps increasing relative to the baseline, but also that low-skill wages are eventually larger than in the baseline case. Although not explicitly modeled here, this suggests that a policy which would aim at helping low-skill workers by taxing automation might temporarily help low-skill workers, but could have negative long-term consequences. A higher productivity for horizontal innovation, γ , implies that *GDP* and low-skill wages initially grow faster than in the baseline. Therefore Phase 2 starts sooner, which explains why the skill premium

jumps relative to the baseline case before increasing smoothly. The asymptotic results can be derived formally (see Appendix 8.3.6), and we establish the following proposition.

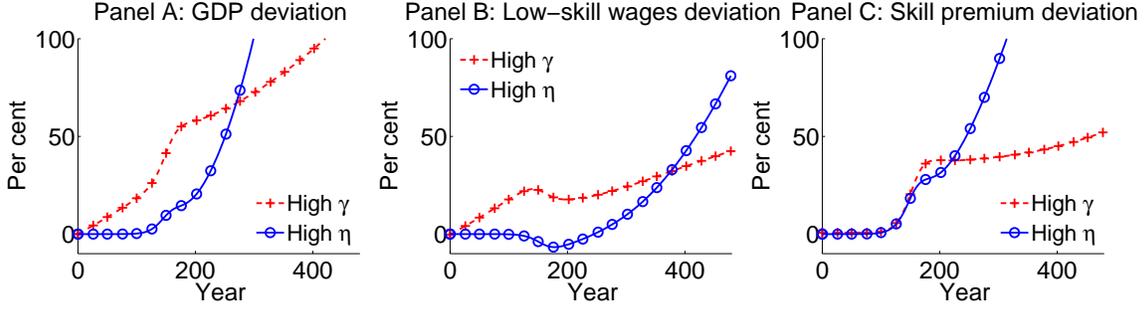


Figure 5: Deviations from baseline model for more productive horizontal innovation technology (γ) and more productive automation technology (η).

Proposition 4. *The asymptotic growth rates of GDP g_{∞}^{GDP} and of low-skill wages g_{∞}^w increase in the productivity of automation η and horizontal innovation γ . The asymptotic share of automated products G_{∞} decreases in γ .*

Growth and automation. For our simulation in Section 3.5, we chose parameters for which the growth rates in Phases 1 and 3 were close, so that Phase 2 coincided with an increase in GDP growth. This need not be the case, and growth in Phase 3 can be either higher or lower than that of Phase 1. Figure 6 shows a case where Phase 3 growth is substantially lower.²⁸ As shown in Panel A, growth is a little higher at the beginning of Phase 2 than in Phase 1 but then it continuously decreases in the second part of Phase 2 and is much lower in Phase 3. Intuitively this occurs because horizontal innovation drops sufficiently during Phase 2 as new non-automated firms find it harder to compete with already automated firms. The lack of an acceleration in GDP growth in recent decades has often been advanced in opposition to the hypothesis that a technological revolution explains the recent increase in the skill-premium (Acemoglu, 2002a). Our model does bear similarities with such a theory (although our technological revolution, automation, can actually be quite progressive and slow); Figure 6 is therefore important in showing that GDP growth need not accelerate. Intuitively, here, the increase in automation research happens at the expense of horizontal innovation.

²⁸The parameters are identical to the baseline case except for: $\sigma = 2.5$, $\beta = 0.55$, $\eta = 0.1$, $\gamma = 0.23$ and $N_0 = 344.25$, these parameters lower the growth rate of the economy particularly in the asymptotic steady-state because automation consumes more resources and is less effective as high-skill workers have a larger factor share in production. N_0 is higher so as to shorten Phase 1 in the graph.

Conversely, making the automation technology more effective (say by reducing the cost share of high-skill workers, β) could create the opposite pattern of a low initial growth rate followed by a higher eventual growth rate. Such might correspond to the transition through the industrial revolution: Before, the industrial revolution the economy is driven by relatively slow horizontal innovation and there is little incentive to innovate. As wages gradually increase, the incentive to automate rises and the engine of economic growth gradually switches from a relatively inefficient horizontal innovation to a more rapid automation innovation with permanently higher growth as a consequence.

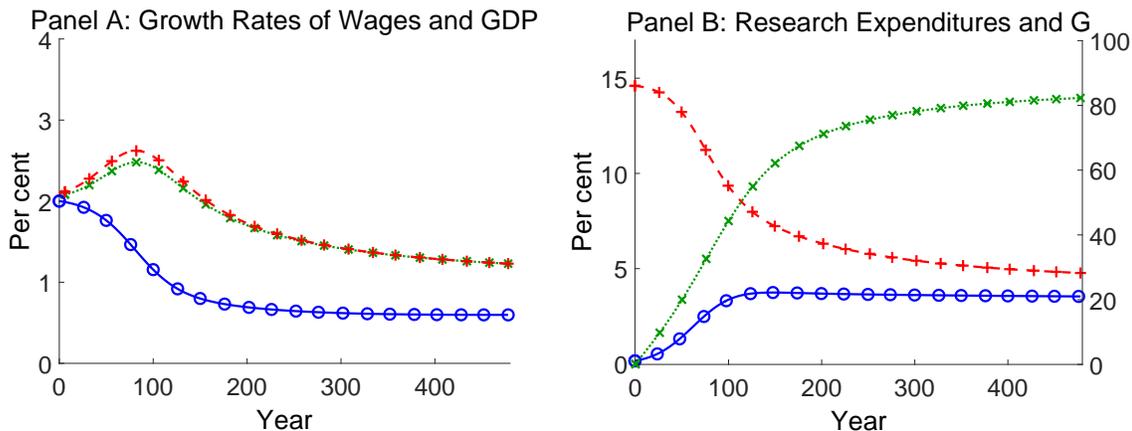


Figure 6: Transitional dynamics with a low growth rate in Phase 2.

3.8 Discussion

Social planner’s problem. The social planner’s problem is studied in Appendix 8.5. The optimal allocation is qualitatively similar to the equilibrium we described, so that our results are not driven by the market structure we imposed. The social planner correct for four market imperfections: a monopoly distortion, a positive externality in horizontal innovation from the total number of products, a positive externality in the automation technology from the total number of products (the term N_t^κ) and a positive externality in the automation technology from the share of automated products when $\tilde{\kappa} > 0$ (which we referred so far as the “automation externality”). The optimal allocation can be decentralized using lump-sum taxes and the appropriate subsidies to the use of intermediates inputs, horizontal innovation and, if $\tilde{\kappa} > 0$, automation.

Comparison with Acemoglu and Restrepo (2015). In ongoing work, Acemoglu and Restrepo (2015) build a model with automation and the creation of new tasks.

Automation plays a similar role in both papers (although in their baseline version, there is only one type of labor), but while in our model new tasks (new products) add up to the stock of existing ones, in their model new tasks are more complex version of existing ones. These new tasks are born non-automated (as in our paper) but with a higher labor productivity (contrary to our paper) which allows them to replace their previous automated vintage. With increasing labor productivity in successive vintages of tasks and under the appropriate assumptions regarding the innovation technology, their model features a stable steady-state.²⁹ The economy self-corrects: if a temporary shock leads to more automation in the short-run, lower wages will reduce incentives to automate, pulling the economy back to its steady-state with balanced growth. Yet, their model cannot explain the origin of such a shock and therefore cannot be used to account for trends such as a secular rise in the skill premium. By contrast our model endogenously explains why automation may become more prevalent as an economy develops.³⁰

In the rest of the paper, we present two extensions of the baseline model: the first one introduces an endogenous supply response in the skill distribution, and is used to perform a quantitative exercise, the second one includes middle-skill workers and allows the model to account for wage polarization. Besides, Appendix 7.3 presents an extension where the production technology for machines and the consumption good differ, and Appendix 7.4 presents an extension where machines are part of a capital stock.

4 Quantitative Exercise

In this section, we conduct a quantitative exercise to compare empirical trends for the United States for the past 50 years with the predictions of our model using Bayesian techniques. As argued in Goldin and Katz (2008), during this period the relative supply of skilled workers increased dramatically so we let workers switch between skill-types in response to changes in factor rewards.

²⁹For this it is crucial that the higher productivity of more complex tasks applies only to labor and not machines. In our model this would be as if in equation (2) there were a labor productivity coefficient $\tau(i)$ in front of $l(i)$ which increases exponentially with i .

³⁰Although this is not the focus of our paper, our model also features elements of self-correction in the presence of exogenous shocks. For instance, in the case with no automation externality ($\bar{\kappa} = 0$), a positive exogenous shock on G_t will be followed by a period where automation is relatively less intense (as the skill premium would have declined), so that eventually the asymptotic share of automated products stays the same.

4.1 An endogenous supply response in the skill distribution

Specifically, let there be a unit mass of heterogeneous individuals, indexed by $j \in [0, 1]$ each endowed with $l\bar{H}$ units of low-skill labor and $\Gamma(j) = \bar{H} \frac{(1+q)}{q} j^{1/q}$ units of high-skill labor (the important assumption here is the existence of a fat tail of individuals with low ability). The parameter $q > 0$ governs the shape of the ability distribution with $q \rightarrow \infty$ implying equal distribution of skills and $q < \infty$ implying a ranking of increasing endowments of high-skill on $[0, \bar{H}(1+q)/q]$. Proposition 3 can be extended to this case and in fact the steady state values $(G^*, \hat{h}^{A*}, g^{N*}, \chi^*)$ are the same as in the model with a fixed high-skill labor supply \bar{H} . Proposition 2 also applies except that the asymptotic growth rate of low-skill wages is higher (see Appendix 8.7):

$$g_\infty^v = g_\infty^Y = \psi g_\infty^N \text{ and } g_\infty^w = \frac{1+q}{1+q+\beta(\sigma-1)} g_\infty^Y. \quad (26)$$

At all points in time there exists an indifferent worker (\bar{j}_t) where $w_t = (1+q)/q(\bar{j}_t)^{1/q} v_t$, with all $j \leq \bar{j}_t$ working as low-skill workers and all $j > \bar{j}_t$ working as high-skill workers. This introduces an endogenous supply response as the diverging wages for low- and high-skill workers encourage shifts from low-skill to high-skill jobs, which then dampens the relative decline in low-skill wages. Hence, besides securing themselves a higher future wage growth, low-skill workers who switch to a high-skill occupation also benefit the remaining low-skill workers. Since all changes in the stock of labor are driven by demand-side effects, wages and employment move in the same direction.

4.2 Bayesian estimation

Because of data availability and to make our exercise easily comparable to the rest of the literature, we focus on the last 50 years. In particular, this allows us to identify low-skill workers with non-college educated workers and high-skill workers with college educated workers. We match the skill-premium and the ratio of skilled to non-skilled workers (both calculated using the methodology of Acemoglu and Autor, 2011) as well as the growth rate of real GDP/employment and the share of labor in total GDP (both taken from the National Income and Products Accounts). We further associate the use of machines with private equipment (real private non-residential equipment, “Table 2.2. Chain-type Quantity Indexes” from NIPA). All time series start in 1963 when the skill-premium and skill-ratio are first available and until 2007 to avoid the Great Recession. We match

the accumulated growth rate of private equipment by indexing both X and real private equipment to 100 in 1963.³¹ We stress that our exercise is much more demanding than previous attempts which feed in input time paths from the data, while we make them endogenous. Both Katz and Murphy (1992) or Golden and Katz (2008) take as given the time paths of labor inputs (endogenous here) and do not attempt to explain the skill bias of technical change which here is constrained to result from economic incentives and decisions. Similarly, Krusell et al. (2000) do not allow for technological change but take the time paths of labor inputs and equipment as given.

Due to the relatively small sample size we use Bayesian techniques to estimate our model, though little would change if we instead employed Maximum Likelihood procedures (in fact since we choose a uniform prior the maximum likelihood point estimate is equal to the mode of the Bayesian estimator). The model presented until now is not inherently stochastic, and in order to bring it to the data, we add normally distributed auto-correlated measurement errors. That is, we consider an economy where the underlying structure is described deterministically by our model, but the econometrician only observes variables with normally distributed auto-correlated measurement errors. With a full parametrization of the model the parameters are not uniquely identified and we restrict $\bar{H} = 1$ without loss of generality. Therefore, our deterministic model has 14 parameters including n_0 and G_0 . Including two parameters (variance and correlation) for each of the five measurement errors, this leaves us with 24 parameters.³² This gives a joint distribution of the observed variables given parameters and, with a chosen prior, standard Bayesian methods can be employed to find the posterior distribution (see Ap-

³¹The use of machines, X , has no natural units and we can therefore not match the level of X . Alternatively, we could normalize X by GDP, but we do not think of equipment as the direct empirical counter-part of X . First, equipment is a stock, whereas X is better thought of as a flow variable. Second many aspects of automation might not be directly captured in equipment. Hence, equipment is better thought of as a proxy for X that grows in proportion to X . Empirically, equipment/GDP is about twice that of our predicted value of X/GDP .

³²More specifically, for time period $t = 1, \dots, T$, let $(Y_1^t, \dots, Y_M^t) \in R^{M \times T}$ be a vector of M predicted variables with time paths of $Y_m^t = (Y_{m,s})_{s=1}^t$ for $m \in \{1, \dots, M\}$ and $Y^t = (Y_1^t, \dots, Y_M^t)$. Let the complete set of parameters in the deterministic model be $b_P \in \mathcal{B}_P \subset R^K$. We can then write the predicted values as $Y_m^T(b_P)$, for $m = 1, \dots, M$. We add normally distributed measurement errors with zero mean to get the predicted values as $\hat{Y}_m^T = Y_m^T + \epsilon_m^T$, where $\epsilon_m^T \sim N(0, \Sigma_m)$ and Σ_m is the covariance matrix of the measurement errors. The errors are independent across types, $E[\epsilon_m^{T'} \epsilon_n^T] = 0$, for $m \neq n$, but potentially auto-correlated: the elements of Σ_m are such that the t, t' -element of Σ_m is given by $\sigma_m^2 \rho_m^{|t-t'|}$, where $\sigma_m^2 > 0$ and $-1 < \rho_m < 1$. Hence, σ_m^2 is the unconditional variance of a measurement error for variable m and ρ_m is its auto-correlation. This gives a total of $2M$ stochastic parameters and we label the combined set of these and b_P as $b \in \mathcal{B} \subset R^{2M+K}$. This leads to a joint probability density for \hat{Y}^T of $f(\hat{Y}^T|b) = \prod_{m=1}^M f_m(\hat{Y}_m^T|b)$ and with a uniform prior $f(b|\hat{Y}^T) \propto f(\hat{Y}^T|b)$.

pendix 8.10 for a full description of the procedure as well as domain on the prior uniform distribution). The domain of the prior is deliberately kept wide for parameters not easily recovered from other studies such as the characteristics of the automation technology.

Table 2 shows the mode of the posterior distribution. The unconditional posterior distribution of each parameter is shown in Figure 22 in Appendix 8.10, which demonstrates that variance for the posterior unconditional distribution is generally small.

Table 2: The mode of the posterior distribution.

	σ	μ	β	l	γ	$\tilde{\kappa}$	θ	η	κ	ρ	φ	q
Mode	4.17	0.66	0.76	0.91	0.20	0.29	2.09	0.27	0.72	0.058	8.60	0.82
	n_0	G_0	ρ_1	σ_1^2	ρ_2	σ_2^2	ρ_3	σ_3^2	ρ_4	σ_4^2	ρ_5	σ_5^2
Mode	0.49	0.59	0.97	0.01	0.99	0.026	0.96	0.001	0.24	0.0004	0.97	0.051

Note: (σ_i^2, ρ_i) , $i = 1, \dots, 5$ estimate refer to skill-premium, skill-ratio, labor share of GDP, growth rate of GDP/employment, and Real Private Equipment, respectively.

Three parameter estimates are worth noting. First, the parameter of the automation externality, $\tilde{\kappa}$, is centered around 0.29 implying a substantial automation externality, a force for an accelerated Phase 2. Second, G_0 is centered around 0.59 implying that Phase 2 was already well underway in the early 1960s. Finally, the estimate of β —the factor share to machines/low-skill workers— of 0.76 implies substantial room for automation.

Figure 7 further shows the predicted path of the matched data series along with their empirical counterparts at the mode of the posterior distribution. Panel A demonstrates that the model matches the rise in the skill-premium from the late 1970s onwards reasonably well, but misses the flat skill-premium in the period before. As argued in Goldin and Katz (2008), the flat skill-premium in this period is best understood as the consequence of a large increase in the stock of college-educated workers caused by other factors than technological change (the Vietnam war and the increase in female college enrollment). Correspondingly, our model, which only allows relative supply to respond to relative factor rewards, fails to capture a substantial increase in the relative stock of skilled labor in the 1960s and early 1970s (Panel B). More importantly, the model predicts a substantially higher drop in the labor share of GDP (14 versus 5 percentage points empirically). The simple structure of the model forces any increase in the use of machines to be reflected in a drop in the labor share. As discussed in Section 3.6, a number of extensions would allow for more flexibility. The model matches the average growth rate of GDP/employment, but as a long-run growth model, is obviously not capable of matching the short-run fluctuations around trend (Panel D).

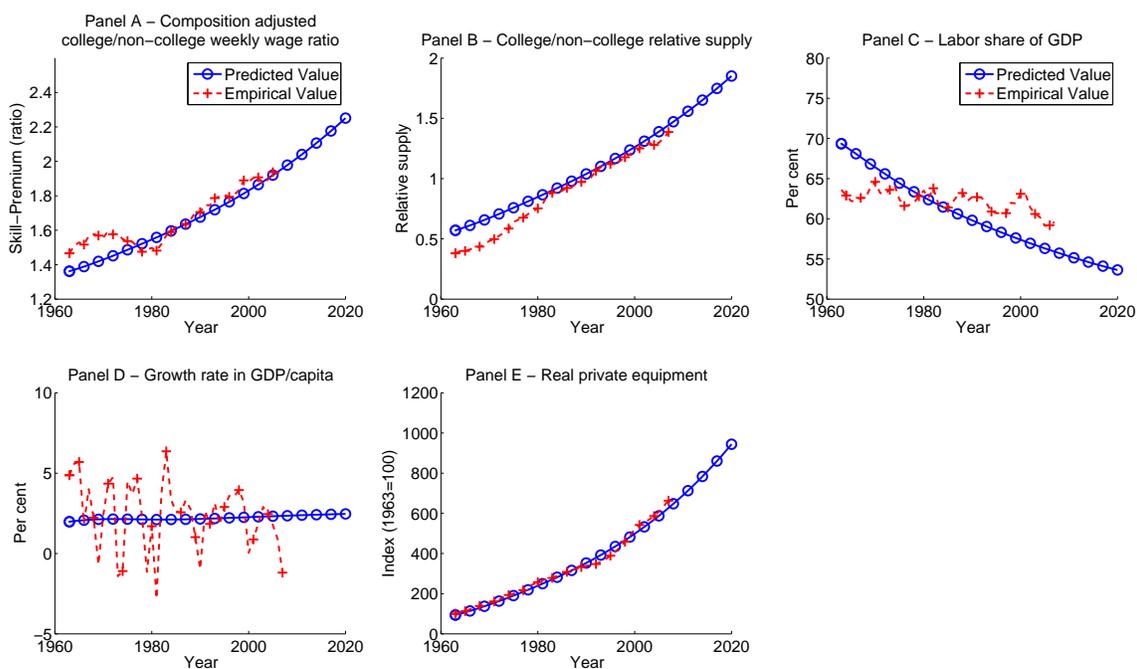


Figure 7: Predicted and Empirical time paths

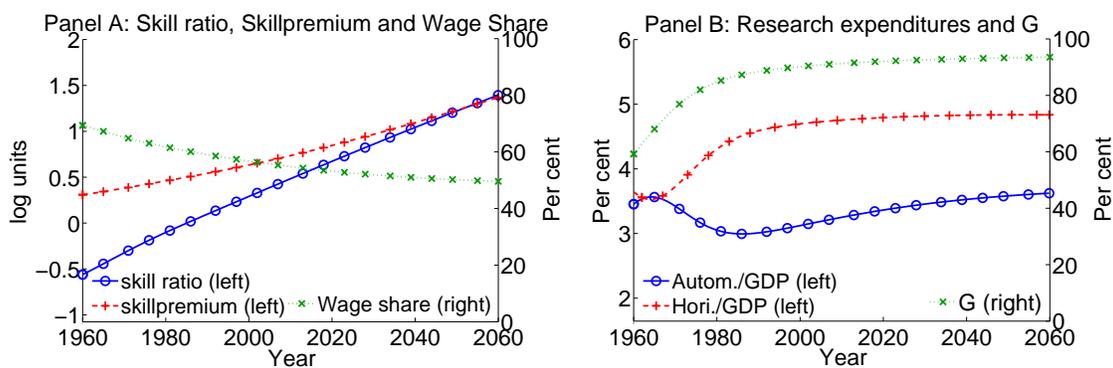


Figure 8: Transitional Dynamics with calibrated parameters

Panel E shows that the model captures the exponential growth in private equipment very well—this is not an automatic consequence of matching the GDP/employment growth rate as equipment has been growing at around 1 percentage point faster than GDP since 1963.

Figure 8 plots the transitional dynamics from 1960 to 2060. Panel A shows that the skill ratio and the skill premium are predicted to keep growing at nearly constant rates, while the labor share is to stabilize at a slightly lower level than today. Panel B suggests

that the share of automated products today is not far from its steady-state value.

5 Middle-Skill Workers and Wage Polarization

As mentioned in the introduction, a recent literature (e.g. Autor et. al., 2006 and Autor and Dorn, 2013) argues that since the 1990s, wage polarization has taken place: inequality has kept rising in the top half of the distribution, but it has narrowed for the lower half. They conjecture that these “middle-skill”-workers are performing cognitive routine tasks which are the most easily automated. Our model suggests a related, but distinct explanation: automating the tasks performed by middle-skill workers is not easier, but more difficult and therefore happened later (or alternatively, the easily automatable low-skill tasks have already been automated). Hence, before 1990 and in fact for most of the 20th century low-skill workers were in the process of being replaced by machines as semi-automated factories, mechanical farming, household appliances etc were increasingly used, whereas since the 1990s, computers are replacing middle-skill workers. In fact, Figure 3 in Autor and Dorn (2013) shows that low-skill workers left non-service occupations from the 70’s, which is consistent with the view that their tasks in non-service occupations were automated before the middle-skill workers’ tasks. As such our model can explain how a phase of wage polarization can follow one of a uniform increase in wage inequality.

To make this precise, we introduce a mass M of middle-skill workers into the model. These workers are sequentially ranked such that high-skill workers can perform all tasks, middle-skill workers can perform middle-skill and low-skill tasks, and low-skill workers can perform only low-skill tasks. All newly introduced intermediate products continue to be non-automated, but there is an exogenous probability δ that they require low-skill and high-skill tasks as described before, and a probability $1 - \delta$ that they require both middle-skill and high-skill tasks in an analogous manner. We refer to the former type of products as “low-skill products” and the latter type as “middle-skill products”. This gives the following production functions (for $i \in [0, N_t]$):

$$y_L(i) = \left[l(i)^{\frac{\epsilon-1}{\epsilon}} + \alpha(i) (\tilde{\varphi}_L x(i))^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon\beta}{\epsilon-1}} h(i)^{1-\beta},$$

$$y_M(i) = \left[m(i)^{\frac{\epsilon-1}{\epsilon}} + \alpha(i) (\tilde{\varphi}_M x(i))^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon\beta}{\epsilon-1}} h(i)^{1-\beta},$$

where $y_L(i)$ and $y_M(i)$ are the production of low-skill and middle-skill products, respectively, and $m(i)$ is the use of middle-skill workers by a firm of the latter type. $\tilde{\varphi}_L$ and $\tilde{\varphi}_M$ are the productivity of machines that replace low-skill and middle-skill workers, respectively. The mass of low-skill products is δN , the mass of middle-skill products is $(1 - \delta)N$ (alternatively all products could be produced by all factors; this would make the analysis substantially more complicated without altering the underlying argument). The final good is still produced competitively by a CES aggregator of all intermediate inputs, and all machines are produced one-for-one with the final good keeping a constant price of 1. The shares of automated products, G_L and G_M will in general differ.

Both types of producers have access to an automation technology as before, but we allow the productivity to differ, such that automation happens with intensity $\eta_L G_L^{\tilde{\kappa}} (Nh_L^A)^{\kappa}$ for low-skill products and $\eta_M G_M^{\tilde{\kappa}} (Nh_M^A)^{\kappa}$ for middle-skill products. The equilibrium is defined analogously to section 3.3 and a proposition analogous to Proposition 3 exists.

We choose $\delta = 1/2$ and set $L = M = 1/3$ and keep parameters as before except that we choose $\tilde{\varphi}_M = 0.15$ and $\tilde{\varphi}_L = 0.3$, to focus on a situation where machines are less productive in middle-skill products than in low-skill ones. The situation would be similar had we chosen $\tilde{\varphi}_M = \tilde{\varphi}_L$, but $\eta_M < \eta_L$ such that the automation technology for middle-skill firms is less productive. Figure 9 describes the equilibrium in the presence of a large externality in the automation technology ($\tilde{\kappa} = 0.5$).

The overall picture is similar to that of Figure 4, but with distinct paths for low-skill and middle-skill wages denoted w and u . One can now distinguish 4 phases. Phase 1 is analogous to Phase 1 in the previous case, and all wages grow at roughly the same rate. From around year 200, low-skill wages become sufficiently high, that low-skill product firms start investing in automation and G_L starts growing. Yet, since machines are less productive in middle-skill workers' tasks, G_M stays low until around year 300. During this second phase, inequality increases uniformly, high-skill wages grow faster than middle-skill wages which again grow faster than low-skill wages. Middle-skill wages do not grow as fast as GDP because automation in low-skill products increases their market share at the expense of the middle-skill products. From around year 300, the economy enters a third phase, where automation in middle-skill products is now intense. As a result, the growth rate of middle-skill wages drops further, such that low-skill wages actually grow faster than middle-skill wages (all along $v_t \geq u_t \geq w_t$, so no group has an incentive to be employed below its skill level).³³ However, depending on parameters,

³³Empirically, the polarization looks more like a J curve than a U curve as the difference in growth

the polarization phase need not be as salient as here (for instance, there is barely any polarization when there is no externality in automation, $\tilde{\kappa} = 0$, but the other parameters are kept identical, see Appendix 7.2.5 for this case).

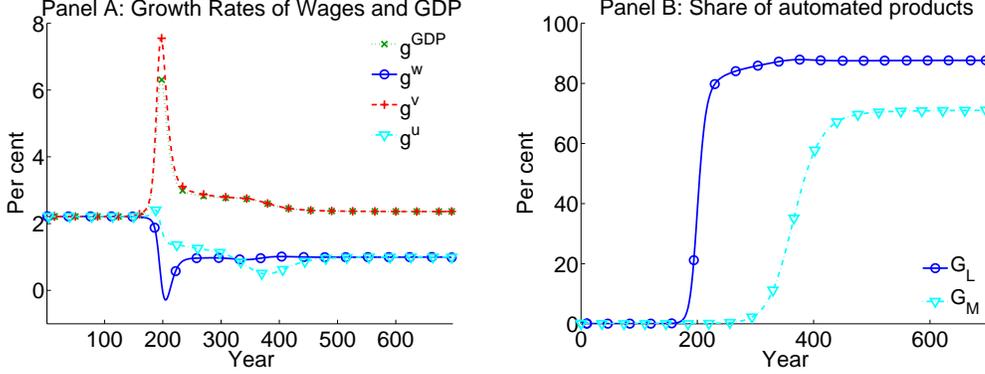


Figure 9: Transitional dynamics with middle-skill workers in the presence of an automation externality ($\tilde{\kappa} = 0.5$).

Finally, in a fourth Phase (from around year 450), G_L and G_M are close to their steady-state levels and the economy approaches the asymptotic steady-state, with low-skill and middle-skill wages growing positively but at a rate lower than that of the economy. Proposition 2 can be extended to this case. High-skill wages and output all grow at the same rate which depends on the growth rate of the number of products, while low-skill and middle-skill wages grow at a lower rate such that:

$$g_\infty^v = g_\infty^Y = \psi g_\infty^N \text{ and } g_\infty^w = g_\infty^u = g_\infty^Y / (1 + \beta(\sigma - 1)).$$

Our model shows how automation may affect more middle-skill workers than low-skill workers, because the economic benefits of automating the former are greater despite a worse automation technology. Yet, some papers argue that the technological opportunities for automation themselves are today lower for low-skill than for middle-skill workers. This is easy to reconcile with our model: assume that for both types of products, a common fixed share can never be automated (for instance because the associated tasks are not routine enough). After the start of Phase 2, the share of low-skill workers hired in products that can never be automated will be larger than the corresponding share for middle-skill workers (since a higher share of low-skill products will already have been automated), so that it will be on average easier to automate a middle-skill product than

rates of wages between the bottom and the middle of the income distribution is modest. Here as well, high-skill wages grow faster than both low-skill and middle-skill wages from the beginning of Phase 2.

a low-skill one. That is, middle-skill workers end up performing on average more routine tasks that are currently being automated (as emphasized by the literature), but this is only the case *because* the routine tasks that were performed by low-skill workers and that could be automated have already been largely automated.

Naturally, the phase of intense automation of middle-skill products may occur sooner than that of low-skill products: for instance if the supply of middle-skill workers is low enough to generate a large middle-skill over low-skill wage ratio. In fact, Katz and Margo (2014) suggest that the recent phase of polarization has a counterpart in the 19th century de-skilling of manufacturing as the tasks of (middle-skilled) artisans got automated, as their wages were much higher than that of unskilled workers (maybe because urbanization increased the supply of low-skill workers).

Automating some high-skill tasks. Alternatively, one may assume that middle-skill and high-skill workers are identical and therefore use this framework to analyze the case where some high-skill production tasks are automatable—in particular, if β is close to (but smaller than) 1, such a model would capture the situation where all production tasks except for headquarter services are automatable. Then, automation will affect in turn both low-skill and high-skill products, with the order depending on the relative wage of both and the relative effectiveness of the two automation technologies. One crucial difference with the middle-skill case is that as some high-skill workers' tasks remain non-automatable (in production and in research), their wage in the long-run still grows at the same rate as *GDP* so that the skill premium keeps rising.

6 Conclusion

In this paper, we introduced automation in a horizontal innovation growth model. We showed that in such a framework, the economy will undertake a structural break. After an initial phase with stable income inequality and stable factor shares, automation picks up. During this second phase, the skill premium increases, low-skill wages stagnate and possibly decline, the labor share drops—all consistent with the US experience in the past 50 years—and growth starts relying increasingly on automation. In a third phase, the share of automated products stabilizes, but the economy still features a constant shift of low-skill employment from recently automated firms to as of yet non-automated firms. With a constant and finite aggregate elasticity of substitution between low-skill workers and machines, low-skill wages grow in the long-run. Wage polarization can be

accounted for once the model is extended to include middle-skill workers.

The model shows that there is a long-run tendency for technical progress to displace substitutable labor (this is a point made by Ray, 2014, in a critique of Piketty, 2014), but this only occurs if the wages of the workers which can be substituted for are large relative to the price of machines. This in turn can only happen under three scenarios: either automation must itself increase the wages of these workers (the scale effect dominates the substitution effect), or there is another source of technological progress (here, horizontal innovation), or technological progress allows a reduction in the price of machines relative to the consumption good (here, only present in Appendix 8.8). Importantly, when machines are produced with a technology similar to the consumption good, automation can only reduce wages temporarily: a prolonged drop in wages would end the incentives to automate in the first place.

Fundamentally, the economy in our model undertakes an endogenous structural change when low-skill wages become sufficiently high. This distinguishes our paper from most of the literature, which seeks to explain changes in the distribution of income inequality through exogenous changes: an exogenous increase in the stock of equipment as per Krusell et al. (2000), a change in the relative supply of skills, as per Acemoglu (1998), or the arrival of a general purpose technology as in the associated literature. This makes our paper closer in spirit to the work of Buera and Kaboski (2012), who argue that the increase in income inequality is linked to the increase in the demand for high-skill intensive services, which results from non-homotheticity in consumption.

The present paper is only a first step towards a better understanding of the links between automation, growth and income inequality. In future research, we will extend it to consider policy implications. The simple sensitivity analysis on the automation technology (section 3.7) suggests that capital taxation will have non-trivial implications in this context. Automation and technological development are also intrinsically linked to the international economy. Our framework could be used to study the recent phenomenon of “reshoring”, where US companies that had offshored their low-skill intensive activities to China, now start repatriating their production to the US after having further automated their production process. Finally, our framework could also be used to study the impact of automation along the business cycle: Jaimovich and Siu (2012) argue that the destruction of the “routine” jobs happens during recessions, which raises the question of whether automation is responsible for the recent “jobless recovery”.

References

- Acemoglu, D. (1998). Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *Quarterly Journal of Economics*, 113(4):1055–1089.
- Acemoglu, D. (2002a). Directed Technical Change. *Review of Economic Studies*, 69(4):871–809.
- Acemoglu, D. (2002b). Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature*, 40(1):7–72.
- Acemoglu, D. (2007). Equilibrium Bias of Technology. *Econometrica*, 75(5):1371–1409.
- Acemoglu, D. (2010). When Does Labor Scarcity Encourage Innovation? *Journal of Political Economy*, 118(6):1037–1078.
- Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, 4(B).
- Acemoglu, D. and Guerrieri, V. (2008). Capital Deepening and Nonbalanced Economic Growth. *Journal of Political Economy*, 116(3):467–498.
- Acemoglu, D. and Restrepo, P. (2015). The race between machine and man: Implications of technology for growth, factor shares and employment.
- Aghion, P. and Howitt, P. (1996). Research and Development in the Growth Process. *Journal of Economic Growth*, 1(1):49–73.
- Aghion, P., Howitt, P., and Violante, G. (2002). General Purpose Technology and Wage Inequality. *Journal of Economic Growth*, 7(4):315–345.
- Autor, D. (2013). The “Task Approach” to Labor Markets: An Overview. *Journal for Labour Market Research*, 46(3):185–199.
- Autor, D. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review*, 103(5):1553–1597.
- Autor, D., Katz, L., and Kearney, M. (2006). The Polarization of the U.S. Labor Market. *American Economic Review*, 96(2):189–194.
- Autor, D., Katz, L., and Kearney, M. (2008). Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics*, 90(2):300–323.
- Autor, D., Katz, L., and Krueger, A. (1998). Computing Inequality: Have Computers Changed the Labor Market? *Quarterly Journal of Economics*, 113(4):1169–1213.
- Autor, D., Levy, F., and Murnane, R. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.
- Bartel, A., Ichniowski, C., and Shaw, K. (2007). How Does Information Technology

- Really Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills. *Quarterly Journal of Economics*, 122(4):1721–1758.
- Benzell, S., Kotlikoff, L., LaGarda, G., and Sachs, J. (2015). Robots are us: some economics of human replacement. NBER wp 20941.
- Brynjolfsson, E. and McAfee, A. (2014). *The Second Machine Age: Work, Progress and Prosperity in a Time of Brilliant Technologies*. W.W. Norton & Company.
- Buera, F. and Kaboski, J. (2012). The Rise of the Service Economy. *American Economic Review*, 102(6):2540–2569.
- Burstein, A., Morales, E., and Vogel, J. (2014). Accounting for Changes in Between-Group Inequality. Working paper.
- Card, D. and DiNardo, J. (2002). Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles. *Journal of Labor Economics*, 20(4):733–783.
- Caselli, F. (1999). Technological Revolutions. *American Economic Review*, 89(1):78–102.
- Costinot, A. and Vogel, J. (2010). Matching and Inequality in the World Economy. *Journal of Political Economy*, 118(4):747–786.
- Elsby, M., Hobijn, B., and Sahin, A. (2013). The Decline of the U.S. Labor Share. *Brookings Papers on Economic Activity*, 47(2):1–63.
- Feng, A. and Graetz, G. (2015). Rise of the machines: The effects of labor-saving innovations on jobs and wages. *Working Paper*.
- Galor, O. and Moav, O. (2000). Ability-Biased Technological Transition, Wage Inequality, and Economic Growth. *Quarterly Journal of Economics*, 115(2):469–497.
- Goldin, C. and Katz, L. (2008). *The Race between Education and Technology*. Harvard University Press.
- Goos, M., Manning, A., and Salomons, A. (2009). Job Polarization in Europe. *American Economic Review*, 99(2):58–63.
- Habakkuk, J. (1962). *American and British Technology in the Nineteenth Century*. Cambridge University Press.
- Hornbeck, R. and Naidu, S. (2014). When the levee breaks: Black migration and economic development in the american south. *American Economic Review*, 104(3):963–990.
- Hornstein, A., Krusell, P., and Violante, G. (2005). The Effects of Technical Change on Labor Market Inequalities. *Handbook of Economic Growth*, 1.
- Jaimovich, N. and Siu, H. (2012). The Trend is the Cycle: Job Polarization and Jobless Recoveries. Working Paper 18334, National Bureau of Economic Research.

- Karabarbounis, L. and Neiman, B. (2014). The Global Decline of the Labor Share. *Quarterly Journal of Economics*, 129(1):61–103.
- Katz, L. and Margo, R. (2014). Technical change and the relative demand for skilled labor: The United States in historical perspective. In Boustan, L., Frydman, C., and Margo, R., editors, *Human Capital in History*, pages 15–57. University of Chicago and NBER.
- Katz, L. and Murphy, K. (1992). Changes in Relative Wages, 1963–1987: Supply and Demand Factors. *Quarterly Journal of Economics*, 107(1):35–78.
- Krusell, P., Ohanian, L., Ríos-Rull, J.-V., and Violante, G. (2000). Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, 68(5):1029–1053.
- Lewis, E. (2011). Immigration, skill mix and capital skill complementarity. *Quarterly Journal of Economics*, 126 (2):1029–1069.
- Lloyd-Ellis, H. (1999). Endogenous Technological Change and Wage Inequality. *American Economic Review*, 89(1):47–77.
- Nelson, R. and Phelps, E. (1966). Investment in Humans, Technological Diffusion, and Economic Growth. *American Economic Review*, 56(1/2):69–75.
- Nordhaus, W. (2007). Two Centuries of Productivity Growth in Computing. *Journal of Economic History*, 67(1):128–159.
- Peretto, P. and Seater, J. (2013). Factor-Eliminating Technical Change. *Journal of Monetary Economics*, 60(4):459–473.
- Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.
- Piketty, T. and Zucman, G. (2014). Capital is Back: Wealth-Income Ratios in Rich Countries 1700–2010. *Quarterly Journal of Economics*, 129(3):1255–1310.
- Ray, D. (2014). Nit-Piketty: A comment on Thomas Piketty’s Capital in the Twenty First Century. Working paper, New York University.
- Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5):71–S102.
- Sachs, J. and Kotlikoff, L. (2012). Smart machines and long-term misery. NBER wp 18629.
- Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, 24(2):235–270.
- Trimborn, T., Koch, K.-J., and Steger, T. (2008). Multi-Dimensional Transitional Dynamics: A Simple Numerical Procedure. *Macroeconomic Dynamics*, 12(3):301–319.

Zeira, J. (1998). Workers, Machines, and Economic Growth. *Quarterly Journal of Economics*, 113(4):1091–1117.