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

David Hémous and Morten Olsen *

June 2020 (first draft: September 2013)

Abstract

We build an endogenous growth model with automation (the replacement of low-skill workers with machines) and horizontal innovation (the creation of new products). Over time, the share of automation innovations endogenously increases through an increase in low-skill wages, leading to an increase in the skill premium and a decline in the labor share. We calibrate the model to the US economy and show that it quantitatively replicates the paths of the skill premium, the labor share and labor productivity. Our model offers a new perspective on recent trends in the income distribution by showing that they can be explained endogenously.

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

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

*David Hémous, University of Zurich and CEPR, david.hemous@econ.uzh.ch, Morten Olsen, University of Copenhagen, mgo@econ.ku.dk. 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 the editor Richard Rogerson and three anonymous referees for their suggestions. We thank Daron Acemoglu, Philippe Aghion, Ufuk Akcigit, Pol Antràs, Tobias Broer, Steve Cicala, Per Krusell, Brent Neiman, Jennifer Page, Ofer Setty, Andrei Shleifer, Che-Lin Su, Fabrizio Zilibotti and Joachim Voth among others for helpful comments. 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, Brown University, Boston University, Yale University, Collège de France, Washington University and the CEPR conference on Growth and Inequality. We thank Ria Ivandic and Marton Varga for excellent research assistance.

1 Introduction

In the past 50 years, the United States has seen dramatic changes in the income distribution. The skill premium increased by 33% between 1963 and 2012 and the labor share has declined by 7 p.p. since the 1970s (Figure 1.A and B). Meanwhile, several automation technologies (numerically controlled machine tools, automatic conveyor systems, industrial robots,...) have been introduced thereby increasing the range of tasks for which machines can substitute for labor. This is supported by patent data which suggest that the share of automation innovation has increased over time (Figure 1.C plots the ratio of automation to non automation patents in machinery in the US according to Dechezleprêtre, Hémous, Olsen and Zanella, 2019).

Our goal is to assess whether these trends can be explained endogenously as reflecting the transitional dynamics of an economy. To do so, we build a model with high- and low-skill workers which combines horizontal innovation (the creation of new products or tasks) and automation. Automation takes place in existing product lines and enables the replacement of low-skill workers with machines. Therefore, our model embodies a task framework where machines can substitute for workers as Autor, Levy and Murnane (2003), directed technical change as Acemoglu (1998) since innovation endogenously occurs in two different technologies, and capital-skill complementarity as Krusell, Ohanian, Ríos-Rull and Violante (2002, henceforth KORV). While these papers rely on exogenous shocks (the advent of computers, an increase in the skill supply prompting a change in the direction of innovation, and a drop in the equipment price, respectively) to explain trends in the income distribution, we argue instead that this can be the result of an *endogenous* increase in the share of automation innovations. Moreover, the interplay between automation and horizontal innovation allows us to account for two puzzles in the literature: the stagnation of labor productivity growth despite the rise in automation innovations and the deceleration of the skill premium since the mid 1990s without an apparent decline in skill-biased technical change (SBTC).

We develop our analysis in three steps. First, we present a version of the model in which technical change is exogenous. Horizontal innovation increases both low-skill and high-skill wages. Within a firm, automation increases the demand for high-skill workers but reduces the demand for low-skill workers. At the aggregate level, automation has an ambiguous effect on low-skill wages and, in line with recent trends, it increases the skill premium and reduces the labor share.

Second, we endogenize innovation, which allows us to rationalize the observed in-

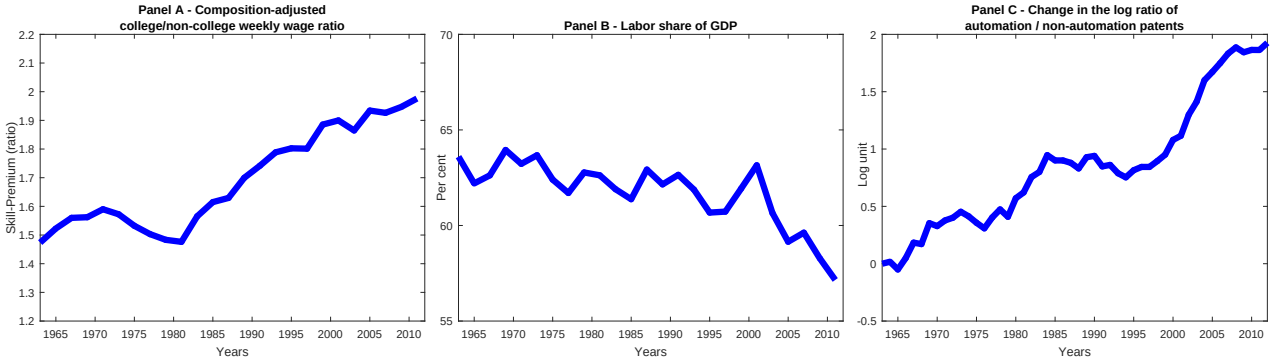

Figure 1: The US skill-premium, labor share and automation innovations. Panel A is taken from Autor (2014). Panel B is from the BLS. Panel C reports the increase in the log ratio of automation to non-automation innovations in machinery in the US according to Dechezleprêtre et al. (2019). See further details in Section 4.

crease in the share of automation innovations. We show that in an economy where low-skill wages are low, there is little automation. As low-skill wages increase with horizontal innovation, the incentive to automate increases and with it the share of automation innovation. As a result, the skill premium rises, the labor share declines and low-skill wages may temporarily decline. Finally, the economy moves toward an asymptotic steady-state where the share of automation innovations stabilizes and low-skill wages grow though slower than high-skill wages and GDP.

In a third step, to assess how far our "endogenous-transition" approach can go quantitatively, we calibrate an extended version of our model to match the evolution of the skill premium, the labor share, productivity and the equipment-to-GDP ratio from 1963 to 2012. Our model captures the trends in the data fairly well. In particular, labor productivity growth stagnates as horizontal innovation declines and the skill premium decelerates in the 1990s and 2000s even though innovation is more directed toward automation.[1] Moreover, conditional on our aggregate production function, a model with exogenous technology would not capture trends better.

We model automation as high-skill-biased following a large literature showing that computerization (Autor, Katz and Krueger, 1998, Autor, Levy and Murnane, 2003, and Bartel, Ichniowski and Shaw, 2007) or industrial robots (Graetz and Michaels, 2018, and

[1]Intuitively, this comes from looking at automation as a stock: with a higher share of automated products, there must be more automation innovation to compensate for its depreciation through horizontal innovation. As a result, our model addresses the Card and DiNardo (2002) critique that the slow-down in the skill premium is inconsistent with the SBTC hypothesis.

Acemoglu and Restrepo, 2017b) decrease the relative demand for low-skill labor.[2]

A large macro literature has argued that SBTC can explain the increase in the skill premium since the 1970's. This literature can be divided into three strands. The first emphasizes Nelson and Phelps (1966)'s hypothesis that skilled workers adapt better to technological change (Lloyd-Ellis, 1999, Caselli, 1999, Galor and Moav, 2000, Aghion, Howitt and Violante, 2002, Beaudry, Green and Sand, 2016). While such theories explain transitory increases in inequality, our model features widening inequality. Yet, we borrow the idea of a shift in production technology spreading through the economy.

A second strand emphasizes the role of capital-skill complementarity: KORV find that the observed increase in the stock of capital equipment can account for most of the variation in the skill premium. Our model also features capital-skill complementarity but differs in several dimensions: it includes low-skill labor-saving innovations; our quantitative exercise is more demanding because we endogenize technology; and we match a decline in the labor share whereas they have a small increase.

A third branch assumes that technical change is either low- or high- skill labor augmenting and measures the bias of technology (Katz and Murphy, 1992, Goldin and Katz, 2008, and Katz and Margo, 2014). The directed technical change literature (Acemoglu, 1998, 2002, 2007) then endogenizes this bias with the skill supply. Such models have no role for labor-replacing technology and cannot generate changes in the labor share (see Acemoglu and Autor, 2011). None of these approaches try to explain features of the income distribution through the transitional dynamics of an economy.

The idea that high wages might incentivize labor-saving technical change dates back to Habakkuk (1962).[3] In Zeira (1998), exogenous increases in TFP raise wages and encourage the adoption of a capital-intensive technology, which further raises wages (while automation can reduce wages in our model). Acemoglu (2010) shows that labor scarcity induces labor-saving innovation. Neither paper analyzes labor-saving innovation in a fully dynamic model nor focuses on income inequality. Peretto and Seater (2013) build a dynamic model of automation where wages are constant. To get a more realistic path for wages, we introduce a second type of innovation, namely the creation of new prod-

[2]Autor, Katz, and Kearney (2006, 2008) and Autor and Dorn (2013) relate wage and job polarization to the computer-driven automation of routine tasks often performed by middle-skill workers. We do not distinguish between low- and middle-skill workers as both have often performed tasks which have later on been automated (a previous version of this paper, Hémous and Olsen, 2016, did so). See also Feng and Graetz (2016).

[3]A few recent papers provide empirical evidence for the role of wages on technology adoption: Lewis (2011), Hornbeck and Naidu (2014), Clemens, Lewis and Postel (2018), Manuelli and Seshadri (2014), Acemoglu and Restrepo (2018b) and Dechezleprêtre et al. (2019).

ucts or tasks. In work subsequent to our paper, Acemoglu and Restrepo (2017a) also develop a growth model where technical change involves automation and the creation of new tasks. While in our model all tasks are symmetric (except for whether they are automated), in theirs, new tasks are exogenously born with a higher labor productivity. As a result, their model features a balanced growth path and they focus on the self-correcting elements of the economy after a technological shock, while we focus on accounting for secular trends. Subsequent papers combining automation and horizontal innovations include Martinez (2018), Rahman (2017) and Zeira and Nakamura (2018).[4]

Section 2 describes the baseline model with exogenous technology. Section 3 endogenizes the path of technology and rationalizes the increase in the share of automation innovations. Section 4 calibrates an extended version of the model. Section 5 concludes. The Main Appendix presents the proofs of the propositions and additional exercises on the quantitative model. The (online) Secondary Appendix presents the proofs of additional results, various extensions and details of the calibration exercise.

2 A Baseline Model with Exogenous Innovation

This section presents a model with exogenous technology to study the consequences of automation and horizontal innovation on factor prices. Section 2.3 derives comparative statics results and relates them to the evolution of the US income distribution. Section 2.4 analyzes the asymptotic behavior of wages for general paths of technology.

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 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 $k \in \{H, L\}$. The final

[4]Benzell, Kotlikoff, LaGarda and Sachs (2017), following Sachs and Kotlikoff (2012) build a model where a code-capital stock can substitute for labor, and show that a technological shock which favors the accumulation of code-capital can lead to lower long-run GDP.

good is produced by a competitive industry combining a set of intermediate products, $i \in [0, N_t]$ 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 $y_t(i)$ is the use of intermediate product i at time t and $\sigma > 1$ is the elasticity of substitution between these products. As in Romer (1990), an increase in N_t represents a source of technological progress.

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 product i is:

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

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

Each product is produced by a monopolist who owns the perpetual rights of production. Production occurs by combining low-skill labor, $l(i)$, high-skill labor, $h(i)$, and, possibly, type-i machines, $x(i)$, according to:

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

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.[5] 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 and a low-skill factor share of β. If it is automated ($\alpha(i) = 1$) machines can be used in the production process as a substitute for low-skill labor with an elasticity $\epsilon > 1$. and we denote The parameter $\tilde{\varphi}$ is the relative productivity advantage of machines over low-skill workers and G denotes the share of automated products. Therefore automation takes the form of a secondary innovation in existing product lines.[6]

Since each product is produced by a single firm, we identify each product with its firm and refer to a firm which uses an automated production process as an automated

[5] We allow for perfect substitutability ($\epsilon = \infty$) in which case $y(i) = [l(i) + \alpha(i)\tilde{\varphi} x(i)]^{\beta} h(i)^{1-\beta}$.

[6] Secondary innovations in a growth model were introduced by Aghion and Howitt (1996) who study the interplay between applied and fundamental research.

firm. We refer to the specific labor inputs provided by high-skill and low-skill workers in the production of different products 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 workers in new tasks. A model with a fixed $\alpha(i)$ for each product would only allow for machines to be used more intensively in production, but always for the same tasks.

Although we will refer to x as "machines", our interpretation also includes any form of computer inputs, algorithms, the services of cloud-providers, etc. In Section 4, we will identify machines with equipment (excluding transport) and software. In turn, automation innovations refer to innovations which allow machines to accomplish tasks with less need for a human operator. This includes robotics but also computer numerical control machine tools, automatic conveyor belts, computer-aided design, etc.[7]

For now, machines are an intermediate input—this assumption is innocuous and in Section 4 machines are a capital input without changing our results qualitatively. Once invented, machines of type i are produced competitively one for one with the final good, such that the price of an existing machine is always equal to 1 and technological progress in machine production follows that in the rest of the economy. Yet, our model can capture the notion of a decline in the real cost of equipment, as automation for firm i can equivalently be interpreted as a decline of the price of machine i from infinity to 1.

2.2 Equilibrium wages

In this section we derive how wages are determined in equilibrium, taking as given the number of products N, the share of automated products G and the employment of high-skill workers in production $H^P \equiv \int_0^N h(i)di$ (we let $H^P \leq H$ to accommodate later sections where high-skill labor is used to innovate).

From equation (2), the unit cost of product i is given by

$$c(w_L, w_H, \alpha(i)) = \beta^{-\beta}(1-\beta)^{-(1-\beta)}(w_L^{1-\epsilon} + \varphi\alpha(i))^{\frac{\beta}{1-\epsilon}} w_H^{1-\beta}, \tag{3}$$

where $\varphi \equiv \tilde{\varphi}^{\epsilon}$, w_L denotes low-skill wages and w_H high-skill wages. For all $w_L, w_H > 0$, automation reduces costs ($c(w_L, w_H, 1) < c(w_L, w_H, 0)$). Price is set as a markup over costs: $p(i) = \sigma/(\sigma-1) \cdot c(w_L, w_H, \alpha(i))$. Using Shepard's lemma and equations (1) and

[7] We include IT innovations in our interpretation because our model does not distinguish between low-skill and middle-skill workers.

(3) delivers the demand for low-skill labor of a single firm.

$$l(w_L, w_H, \alpha(i)) = \beta \frac{w_L^{-\epsilon}}{w_L^{1-\epsilon} + \varphi\alpha(i)} \left(\frac{\sigma - 1}{\sigma}\right)^{\sigma} c(w_L, w_H, \alpha(i))^{1-\sigma} Y. \tag{4}$$

The effect of automation on demand for low-skill labor in a firm is generally ambiguous. This is due to the combination of a negative *substitution* effect (automation allows for substitution between machines and low-skill workers) and a positive *scale* effect (automation decreases costs, lowers prices and increases production). As we focus on labor-saving innovation, we impose the condition $\epsilon > 1 + \beta(\sigma - 1)$ *throughout* the paper which is necessary and sufficient for the substitution effect to dominate at the firm level and ensures $l(w_L, w_H, 1) < l(w_L, w_H, 0)$ for all $w_L, w_H > 0$.

At the aggregate level, since both automated firms and non-automated firms are symmetric, output can written as:

$$Y = N^{\frac{1}{\sigma-1}} \times \left((1-G)^{\frac{1}{\sigma}} \left(\underbrace{(L^{NA})^{\beta} (H^{P,NA})^{1-\beta}}_{T_1} \right)^{\frac{\sigma-1}{\sigma}} + G^{\frac{1}{\sigma}} \left(\underbrace{[(L^A)^{\frac{\epsilon-1}{\epsilon}} + (\tilde{\varphi} X)^{\frac{\epsilon-1}{\epsilon}}]^{\frac{\epsilon\beta}{\epsilon-1}} (H^{P,A})^{1-\beta}}_{T_2} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{5}$$

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 first term T_1 captures the classic case where production takes place with constant shares between factors (low-skill and high-skill labor). The second term T_2 represents the factors used within automated products and features 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.[8]

With CRS and perfect competition in final good production, the price of the final good is equal to its cost. Using that all intermediate producers charge the same mark-up $\sigma/(\sigma - 1)$ and that final good output obeys equation (5), the price normalization gives:

$$\frac{\sigma}{\sigma - 1} \frac{N^{\frac{1}{1-\sigma}}}{\beta^{\beta}(1-\beta)^{1-\beta}} \left(G\left(\varphi + w_L^{1-\epsilon}\right)^{\mu} + (1-G) w_L^{\beta(1-\sigma)} \right)^{\frac{1}{1-\sigma}} w_H^{1-\beta} = 1, \tag{6}$$

[8]For automated firms, this model features an elasticity of substitution between high-skill labor and machines equal to that between high-skill and low-skill labor. This, however, does not hold at the aggregate level, consistent with KORV, who argue that the aggregate elasticity of substitution between high-skill and low-skill labor is greater than that between high-skill labor and machines.

where we define $\mu \equiv \beta(\sigma-1)/(\epsilon-1) < 1$ (by our assumption on ϵ). This relationship defines the *unit isocost curve* in the (w_L, w_H) space in Figure 2. It shows the positive relationship between real wages and the level of technology given by N, the number of products, and G the share of automated firms.

Applying Shepard's lemma to the cost function defined by the left-hand side of (6), we get the relative labor shares in production:[9]

$$\frac{w_H H^P}{w_L L} = \frac{1-\beta}{\beta} \frac{G + (1-G)(1+\varphi w_L^{\epsilon-1})^{-\mu}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{-1} + (1-G)(1+\varphi w_L^{\epsilon-1})^{-\mu}}. \tag{7}$$

This expression gives the *relative demand curve* for high-skill and low-skill labor drawn in Figure 2. Together (6) and (7) determine real wages uniquely as a function of N, G and H^P. For $G = 0$, the relative demand curve is a straight line, with slope $(1-\beta)L/(\beta H^P)$, reflecting the constant factor shares in a Cobb-Douglas economy. For $G > 0$, the right-hand side of (7) increases in w_L, so that the relative demand curve is non-homothetic and bends counter-clockwise as w_L grows. Therefore, as long as G tends toward a positive constant, low-skill and high-skill wages cannot grow at the same rate in the long-run.

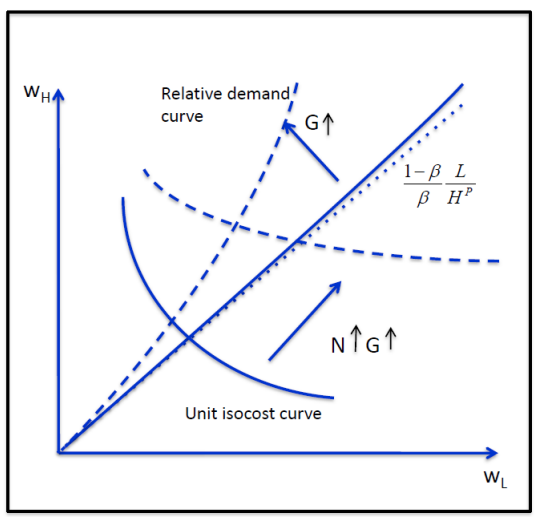

Figure 2: Relative demand curve and isocost curve for different values of N and G.

Intuitively, higher low-skill wages increase the ratio of high-skill to low-skill labor share in production for two reasons. First, they induce more substitution toward machines in automated firms as their use relative to low-skill labor obeys $x/l = \varphi w_L^{\epsilon}$,

[9] When $\epsilon = \infty$, the skill premium is given by $\frac{w_H}{w_L} = \frac{1-\beta}{\beta}\frac{L}{H^P}$ if $w_L < \widetilde{\varphi}^{-1}$ such that no firm uses machines, and $\frac{w_H}{w_L} = \frac{1-\beta}{\beta}\frac{L}{H^P}\frac{G+(1-G)(\widetilde{\varphi}w_L)^{-\beta(\sigma-1)}}{(1-G)(\widetilde{\varphi}w_L)^{-\beta(\sigma-1)}}$ if $w_L > \widetilde{\varphi}^{-1}$.

reflected by the term $(1+\varphi w_L^{\epsilon-1})^{-1}$ in (7)—recall that $\epsilon > 1$. Second, higher low-skill wages improve the cost-advantage of automated firms and their market share. Using (1) and (3), the relative revenues (and profits) of non-automated and automated firms are:

$$R(w_L, w_H, 0)/R(w_L, w_H, 1) = \pi(w_L, w_H, 0)/\pi(w_L, w_H, 1) = (1+\varphi w_L^{\epsilon-1})^{-\mu}, \qquad (8)$$

which decreases in w_L (reflected by the term $(1+\varphi w_L^{\epsilon-1})^{-\mu}$ in (7)).

The static equilibrium is closed by the final good market clearing condition $Y = C + X$, where $C = C_L + C_H$ is total consumption. GDP includes the payment to labor and aggregate profits, which are a share $1/\sigma$ of output. Therefore GDP and the total labor share LS are given by:

$$GDP \equiv \frac{1}{\sigma}Y + w_L L + w_H H, \quad LS = 1 - \frac{1}{1+(\sigma-1)(1-\beta)\left(\frac{w_L L}{w_H H^P} + \frac{H}{H^P}\right)}, \qquad (9)$$

where the second equality uses that payment to high-skill labor in production is a constant share $(1-\beta)(\sigma-1)/\sigma$ of output.

2.3 Technical change and wages

We analyze the consequences of technical change on the level of wages using Figure 2. An increase in the number of products, N, pushes out the isocost curve and increases both low-skill and high-skill wages. When $G = 0$, both wages grow at the same rate since the relative demand curve is a straight line, but for $G > 0$, the demand curve is non-homothetic and the skill premium grows. Therefore, an increase in N at *constant* $G\,(> 0)$ is high-skill biased.

An increase in the share of automated products G has a positive effect on high-skill wages and the skill premium but an ambiguous effect on low-skill wages: Higher automation increases the productive capability of the economy and pushes out the isocost curve (an *aggregate scale* effect), which increases low-skill wages. Yet, it also allows for easier substitution away from low-skill labor which pivots the relative demand curve counter-clockwise (an *aggregate substitution* effect), decreasing low-skill wages. Therefore automation is always high-skill labor biased (w_H/w_L increases) but low-skill labor saving (w_L decreases) if and only if the aggregate substitution effect dominates the aggregate scale effect. Formally, one can show (proof in Appendix 6.1):[10]

[10] In the perfect substitute case, $\epsilon = \infty$, w_H increases in N and weakly increases in G, w_H/w_L

Proposition 1. *Consider the equilibrium* (w_L, w_H) *determined by equations (6) and (7). Assume that* $\epsilon < \infty$*. 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* (w_H) *and low-skill wages* (w_L)*. Provided that* $G > 0$*, an increase in* N *also increases the skill premium* w_H/w_L *and decreases the labor share.*

B) An increase in the share of automated products G *(keeping* N *and* H^P *constant) increases the high-skill wages* w_H*, the skill premium* w_H/w_L *and decreases the labor share. Its impact on low-skill wages is generally ambiguous, but low-skill wages are decreasing in* G *if i)* $1 \leq (\sigma - 1)(1 - \beta)$ *or if ii)* N *and* G *are high enough.*

C) An increase in the number of non-automated products (an increase in N *keeping* GN *constant) increases both high-skill* (w_H) *and low-skill wages* (w_L)*. If* N *is large enough or* $\epsilon < \sigma$*, it decreases the skill-premium.*

Part B gives sufficient conditions under which automation is low-skill labor saving. The aggregate substitution effect is larger than the scale effect in two cases: *i*) The elasticity of substitution σ is large, as newly automated products gain a larger market share; or the cost share of the low-skill labor-machines aggregate β is small as the cost-saving effect of automation is small. *ii*) G and N are large: in that case additional automation hurts low-skill workers more as there are few non-automated firms, while most of the aggregate productivity gains are already realized. It is worth comparing the effect of automation with that of an increase in machines' productivity φ (equivalent to a decline in the price of machine). The latter also has an ambiguous effect on low-skill wages resulting from the combination of a substitution effect and a scale effect, but it is less likely to be low-skill labor saving than automation.[11]

Part C considers an increase in the number of non-automated products, which corresponds to the "horizontal innovation" to be introduced in Section 3. Such technological change pushes out the iso-cost curve (N increases) but also makes the relative demand curve rotate clockwise (GN stays constant). This increases demand for both types of workers and therefore both wages. Horizontal innovation is low-skill labor biased (it re-

weakly increases in N and G and w_L weakly increases in N and weakly decreases in G provided that $1/(1-\beta) \leq \sigma - 1$ or G is large enough. When $\epsilon = \infty$ and $G = 1$, the isocost curve has a horizontal arm and the relative demand curve a vertical one.

[11]Formally, we show that $\frac{\partial w}{\partial \varphi} < 0$ implies that $\frac{\partial w}{\partial G} < 0$ but the reverse is not true–see Appendix 6.1. Intuitively, this is the case because an increase in automation not only acts as "factor augmenting technical change" for the inputs within automated firms, but also as "factor-depleting technical change" for the inputs in non-automated firms. This point can be seen from equation (5) and is made by Aghion, Jones and Jones (2017).

duces the skill premium) if N is large in which case w_L is large so that the isocost curve does not move much with horizontal innovation, or if machines and low-skill workers are not too substitute ($\epsilon \leq \sigma$ is a sufficient condition).[12]

Proposition 1 offers important insights into the types of technological change that can simultaneously account for a rising skill premium (Stylized fact 1) and a decreasing labor share (Stylized fact 2). Specifically, these trends are consistent with an economy with a growing number of products and a constant or rising share of automated products.[13] Moreover, it suggests that an increase in automation may also give rise to a decrease in wages for low-skill workers. Section 3 will model innovation and explain why we should expect a rising number of products and a rising share of automated products until it approaches an asymptotic steady-state value.

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 $g_\infty^a = lim_{t\to\infty} g_t^a$ if it exists. We focus here on the case where the share of automated products admits an interior limit $G_\infty \in (0,1)$ for which we obtain (proof in Appendix 6.2, Appendix 7.2.1 studies the cases where $G_\infty = 0$ or 1):

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 limits G_∞, g_∞^N and H_∞^P with $G_\infty \in (0,1)$, $g_\infty^N > 0$ and $H_\infty^P > 0$. Then, the asymptotic growth of high-skill wages w_{Ht} and output Y_t are:*

$$g_\infty^{w_H} = g_\infty^Y = g_\infty^N / \left((1-\beta)(\sigma-1)\right), \tag{10}$$

and the asymptotic growth rate of w_{Lt} is given by

$$g_\infty^{w_L} = g_\infty^Y / \left(1+\beta(\sigma-1)\right). \tag{11}$$

This proposition first relates the growth rates of output and high-skill wages to

[12] In the perfect substitute case, an increase in the number of non-automated products increases w_H and weakly increases w_L. If $G < 1$ and N is large enough, it decreases the skill premium.

[13] The decreasing relationship between the labor share and the skill premium obtained when machines are an intermediate input generalizes to the case where machines are part of a capital stock with a perfectly elastic supply (see Section 4 and Appendix 7.10).

the growth rate of the number of products. Y_t is proportional to $N_t^{1/((1-\beta)(\sigma-1))}$. This reflects the standard expanding-variety model gains and, in the presence of automation ($G_\infty > 0$), a multiplier effect which increases in the asymptotic share of machines β as machines are a reproducible input.

Second, when there is a positive share of non-automated products asymptotically, $G_\infty < 1$, low-skill workers and machines are imperfect substitutes in the aggregate (even if there are perfect substitute at the firm level $\epsilon = \infty$).[14] As a result, low-skill wages must grow at a positive rate asymptotically when the number of products grows. Intuitively, a growing stock of machines and a fixed supply of low-skill labor imply that the relative price of a worker (w_{Lt}) 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 equals 1 and the real wage w_{Lt} must grow at a positive rate.

Third, the proposition shows that if $G_\infty > 0$, low-skill wages cannot grow at the same rate as output. This results follows from Uzawa's theorem: equation (5) shows that an increase in the number of product N_t is not labor-augmenting unless $G_\infty = 0$. We get $G_\infty > 0$ as long as the automation intensity is bounded away from 0 (see Appendix 7.2.2). Further, when $G_\infty < 1$, the demand for low-skill labor increasingly comes from the non-automated firms (as automation is labor-saving at the firm level), while most of the demand for high-skill labor comes from automated firms. With growing wages, the relative market share of non-automated firms decreases in proportion to $(1 + \varphi w_{Lt}^{\epsilon-1})^{-\mu} \sim \varphi^{-\mu} w_{Lt}^{-\beta(\sigma-1)}$. Then, the growth rate of low-skill wages is a fraction of the growth rate of high-skill wages given by (11). The ratio between the growth rates of high- and low-skill wages 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. Yet, it is independent of the elasticity of substitution between machines and low-skill workers, ϵ or of the exact asymptotic share of automated products G_∞. In this case, non-automated products provide employment opportunities for low-skill workers which limits the relative losses of low-skill workers compared to high-skill workers (their wages grow according to (11) instead of (69) and $\epsilon > 1 + \beta(\sigma - 1)$). In the model of Section 3, the economy endogenously ends up in this case.

[14] As long as new non-automated products are continuously introduced, and the intensity at which non-automated firms are automated is bounded, the share of non-automated products is always positive, i.e. $G_\infty < 1$ (see proof in Appendix 7.2.2). This ensures that there is no economy-wide perfect substitution between low-skill workers and machines.

Proposition 2 establishes general conditions under which low-skill wages grow aymptotically but slower than high-skill wages. We briefly discuss the robustness of this result. First, one might be concerned that the slower growth in low-skill wages is an artifact of having exogenous supplies of low- and high-skill labor. Appendix 7.3 extends our model to allow for endogenous skill choice. Specifically, we consider a Roy model in which workers have heterogeneous comparative advantage between being low- and high-skill. Low-skill wages still grow slower than high-skill wages asymptotically, and now the share of low-skill workers tends toward zero. Second, the result that imperfect aggregate substitution between machines and low-skill workers leads to positive growth in low-skill wages asymptotically relies on machines and the consumption good sharing the same production technology. Appendix 7.4 relaxes this assumption and allows for negative growth in p_t^x. In that case low-skill wages may but need not decline asymptotically.

3 Endogenous innovation

We now model automation and horizontal innovation as the result of investment. In Sections 3.1-3.3, we analyze the effect of wages on innovation (the reverse of Proposition 1), study the transitional dynamics of the system and explain why the economy should experience an increase in the share of automated products as it develops. Section 3.4 explores the interactions between the two innovation processes. Appendix 7.6 provides numerical examples to illustrate this section and shows comparative statics results.

3.1 Modeling innovation

If a non-automated firm hires $h_t^A(i)$ high-skill workers to perform automation research, it becomes automated at a Poisson rate $\eta G_t^{\widetilde{\kappa}} \left(N_t h_t^A(i)\right)^{\kappa}$. Once a firm is automated it remains so forever. $\eta > 0$ denotes the productivity of the automation technology, $\kappa \in (0,1)$ measures its concavity, $G_t^{\widetilde{\kappa}}$, $\widetilde{\kappa} \in [0, \kappa]$, represents possible knowledge spillovers from the share of automated products, and N_t represents knowledge spillovers from the total number of products. The spillovers in N_t ensure that both automation and horizontal innovation may take place in the long-run; they exactly compensate for the mechanical reduction in the amount of resources for automation available for each product (namely high-skill workers) when the number of product increases.[15] The presence of spillovers

[15]These spillovers can be micro-funded as follows: let there be a fixed mass one of firms indexed by j each producing a continuum N_t of products indexed by i so that production is given by $Y_t =$

in automation technology ($\widetilde{\kappa} > 0$) implies a delayed and faster rise in the share of automated products.[16] We assume that $\widetilde{\kappa} < 1-\kappa(1-\beta)$ which ensures that automation always takes off (see Proposition 3).

New products are developed by high-skill workers in a standard manner according to a linear technology with productivity γN_t. With H_t^D high-skill workers pursuing horizontal innovation, the mass of products evolves according to:

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

We assume that firms do not exist before their product is created and therefore cannot invest in automation. As a result, new products are born non-automated, which means that "horizontal innovation" corresponds to an increase in N_t keeping $G_t N_t$ constant and is low-skill biased under certain conditions (see Proposition 1). This is motivated by the idea that when a task is new and unfamiliar, solving unforeseen problems requires the flexibility and outside experience of human workers. Only as the task becomes routine and potentially codefiable, a machine (or an algorithm) can perform it (Autor, 2013).

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

$$\dot{G}_t = \eta G_t^{\widetilde{\kappa}} \left(N_t h_t^A\right)^{\kappa} (1-G_t) - G_t g_t^N. \tag{12}$$

Therefore, the level of automation in the economy, G_t, can be understood as a "stock" that depreciates through the introduction of new products. As a result, for a given growth rate in the number of products (g_t^N), a higher automation intensity per product ($\eta G_t^{\widetilde{\kappa}} \left(N_t h_t^A\right)^{\kappa} (1-G_t)$) is required to increase the share of automated products when this share G_t is already high. This feature plays a role in explaining why the growth rate of the skill premium need not grow the fastest when innovation is the most directed toward automation. It is also one of the main differences between our modeling of automation

$(\int_0^1 \int_0^{N_t(j)} y_t(i,j)^{\frac{\sigma-1}{\sigma}} di dj)^{\frac{\sigma}{\sigma-1}}$. When a firm hires $\widetilde{H}_t^A(j)$ high-skill workers in automation each of its non-automated products gets independently automated with a Poisson rate of $\eta G_t^{\widetilde{\kappa}}[\widetilde{H}_t^A(j)/(1-G_t(j)))]^{\kappa}$. The aggregate economy would be identical to ours and have the same social planner allocation (the decentralized equilibrium would behave similarly but the externality in the automation technology from the number of products would be internalized).

[16] Whenever $\widetilde{\kappa} > 0$, we assume that $G_0 > 0$. Growth models with more than one type of technology often feature similar knowledge spillovers (e.g. Acemoglu, 2002b). Bloom, Schankerman and Van Reenen (2013) show empirically that technologies which are closer to each other in the technology space have larger knowledge spillovers

and a simple reduction in the price of equipment.

Overall, the rate and direction of innovation depends on the equilibrium allocation of high-skill workers between production, automation and horizontal innovation.[17] We define the total mass of high-skill workers working in automation as $H_t^A \equiv \int_0^{N_t} h_t^A(i)di$.[18] High-skill labor market clearing then leads to

$$H_t^A + H_t^D + H_t^P = H. \tag{13}$$

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 $\pi_t^A \equiv \pi(w_{Lt}, w_{Ht}, 1)$ the profits at time t of an automated firm. The asset pricing equation for an automated firm is given by:

$$r_t V_t^A = \pi_t^A + \dot{V}_t^A. \tag{14}$$

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 invests in automation. Denoting by V_t^N the value of a non-automated firm and letting $\pi_t^N \equiv \pi(w_{Lt}, w_{Ht}, 0)$, we get the asset pricing equation:

$$r_t V_t^N = \pi_t^N + \eta G_t^{\widetilde{\kappa}} \left(N_t h_t^A\right)^{\kappa} \left(V_t^A - V_t^N\right) - w_{Ht} h_t^A + \dot{V}_t^N, \tag{15}$$

where h_t^A is the mass of high-skill workers in automation research hired by a single non-automated firm. This equation is similar to equation (14), but profits are augmented by the instantaneous expected gain from innovation $\eta G_t^{\widetilde{\kappa}} \left(N_t h_t^A\right)^{\kappa} \left(V_t^A - V_t^N\right)$ net of expenditure on automation research, $w_{Ht} h_t^A$. This gives the first order condition:

$$\kappa \eta G_t^{\widetilde{\kappa}} N_t^{\kappa} \left(h_t^A\right)^{\kappa-1} \left(V_t^A - V_t^N\right) = w_{Ht}. \tag{16}$$

h_t^A increases with the difference in value between automated and non-automated firms,

[17] We focus here on the decentralized equilibrium, but Appendix 7.8 studies the social planner's problem. The optimal allocation is qualitatively similar so that our results are not driven by the market structure we impose.

[18] Using that by symmetry the total amount of high-skill workers hired in automation research is $H_t^A = (1 - G_t) N_t h_t^A$, we can rewrite (12) as $\dot{G}_t = \eta G_t^{\widetilde{\kappa}} \left(H_t^A\right)^{\kappa} (1 - G_t)^{1-\kappa} - G_t g_t^N$.

and thereby current and future low-skill wages—all else equal.[19]

Free-entry in horizontal innovation ensures that the value of creating a new firm equals its opportunity cost when there is strictly positive horizontal innovation ($\dot{N}_t > 0$):

$$\gamma N_t V_t^N = w_{Ht}. \tag{17}$$

The low-skill and high-skill representative households' problems are standard and lead to Euler equations which in combination give[20]

$$\dot{C}_t / C_t = (r_t - \rho) / \theta, \tag{18}$$

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:

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

3.3 Description of the dynamic equilibrium

Appendix 6.3 shows that the equilibrium can be characterized by a system of 4 differential equations with two state variables (determining N_t and G_t), two control variables (which give the allocation of high-skill workers in innovation and production) and an auxiliary equation defining low-skill wages. It further establishes:

Proposition 3. *Assume that* $\kappa^{-\kappa} \left(\gamma(1-\kappa)/\rho\right)^{\kappa-1} \rho/\eta + \rho/\gamma < \psi H$. *Then:*

A. The system of differential equations admits an asymptotic steady-state with a constant share of automated products $G_\infty \in (0,1)$, *positive growth in the number of products* $g_\infty^N > 0$. *The growth rates of output and wages are given by Proposition 2 as (10) and (11).*

B. For any $a > 0$, *and any* $\hat{t} < \infty$, *there exists an* N_0 *sufficiently low such that during the interval* $\left(0, \hat{t}\right)$, *the automation rate* $\eta G_t^{\tilde{\kappa}} (\hat{h}_t^A)^\kappa < a$ *and the economy behaves arbitrarily close to that of a Romer model where automation is impossible.*

[19] The model predicts that the ratio of high-skill to low-skill labor *in production* is higher for automated than non-automated firms, though not *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).

[20] Consumption growth is the same for both households even though low-skill wages grow at a lower rate than high-skill wages in the long-run. This is possible because low-skill workers save initially anticipating a drop in labor income.

C. If g_t^N admits a positive limit and G_t admits a limit, then the economy converges to an asymptotic steady-state with $G_\infty \in (0,1)$ as described in A.

Proposition 3 establishes three results. First, under the appropriate parameter conditions, there exists an asymptotic steady-state where the share of automated products G_∞ is between 0 and 1.[21,22] Second, for N_0 sufficiently low, the economy behaves close to a Romer model with no automation. Third, if G_t admits a limit and g_t^N a positive limit, the economy must converge toward the asymptotic steady-state regardless of the initial values (N_0, G_0). Therefore, the economy must feature a period where the rate of automation innovation increases: it is low for low N_0 but must be positive later on to ensure a positive share of automated products in the long-run. In other words, the path of technological change itself will be unbalanced through the transitional dynamics. To understand this result, we now explain the evolution of the automation incentives, which, we show, are crucially linked to the level of low-skill wages.[23]

Following (16), the mass of high-skill workers in automation $(H_t^A = (1-G_t)\, N_t h_t^A)$ and therefore the automation intensity rate, given by $\eta G_t^{\widetilde{\kappa}} \left(H_t^A / (1-G_t)\right)^\kappa$, depends on the ratio between the gain in firm value from automation $V_t^A - V_t^N$, and its effective cost namely the high-skill wage divided by the number of products w_{Ht}/N_t:

$$H_t^A = (1-G_t)\left(\kappa\eta G_t^{\widetilde{\kappa}} \frac{V_t^A - V_t^N}{w_{Ht}/N_t}\right)^{1/(1-\kappa)}. \tag{19}$$

Crucially, as the number of products in the economy increases, the ratio $\left(V_t^A - V_t^N\right) / \left(w_{Ht}/N_t\right)$ evolves. Combining (14), (15) and (16) gives the difference in value between an automated and a non-automated firm:

$$V_t^A - V_t^N = \int_t^\infty \exp\left(-\int_t^\tau r_u du\right)\left(\pi_\tau^A - \pi_\tau^N - \frac{1-\kappa}{\kappa} w_{H\tau} h_\tau^A\right) d\tau, \tag{20}$$

[21] To understand the condition $\kappa^{-\kappa}\left(\gamma(1-\kappa)/\rho\right)^{\kappa-1}\rho/\eta + \rho/\gamma < \psi H$, let the efficiency of the automation technology η be arbitrarily large such that the model approaches a Romer model with only automated firms. Then this inequality becomes $\rho/\gamma < \psi H$, which is the 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 and the condition is more stringent.

[22] "Asymptotic" because the system of differential equations only admits a fixed point for $N_t = \infty$. Technically, there is a steady state for the transformed system in Appendix 6.3.1, where the number of product N_t is replaced by an inversely related variable n_t, which is 0 in steady state. Further, multiple steady-states with $G_\infty > 0$ are technically possible but unlikely for reasonable parameter values (see Appendix 7.5.2). In all our numerical simulations, the steady-state was unique and saddle-path stable.

[23] These results extend to the case where the skill supply is endogenous. See Appendix 7.9.

which is the discounted difference of profit flows adjusted for the cost and probability of automation. Recall that Cobb-Douglas production and isoelastic demand imply that both high-skill wages (for given H_t^P) and aggregate profits are proportional to aggregate output. Therefore, w_{Ht}/N_t is proportional to average profits: $w_{Ht}/N_t = \left[G_t\pi_t^A + (1-G_t)\,\pi_t^N\right] / \left[\psi H_t^P\right]$. As a result, the mass of high-skill workers in automation moves with the discounted flow of profits of automated versus non-automated firms divided by average firm profits. Intuitively, with a positive discount rate, $V_t^A - V_t^N$ moves like $\pi_t^A - \pi_t^N$ as a first approximation (from equation (20)), so that we get:

$$\frac{V_t^A - V_t^N}{w_{Ht}/N_t} \tilde{\propto} \frac{\pi_t^A - \pi_t^N}{G_t\pi_t^A + (1-G_t)\,\pi_t^N} = \frac{1-\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-\mu}}{G_t + (1-G_t)\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-\mu}}, \tag{21}$$

where $\tilde{\propto}$ denotes "approximately proportional" and we used equation (8). This highlights low-skill wages (relative to the inverse productivity of machines $\tilde{\varphi}^{-1}$) as the key determinant of automation innovations. When $w_{Lt} \approx 0$ the incentive for automation innovation is very low, whereas when $w_{Lt} \to \infty$ it approaches a positive constant. This price effect bears similarity to Zeira (1998), where the adoption of a labor-saving technology also depends on the price of labor.[24]

Low automation. When the number of products, N_t, is low enough that w_{Lt} is small relative to $\tilde{\varphi}^{-1}$, the difference in profits between automated and non-automated firms is small relative to average profits. Following (19) and (21), the allocation of high-skill labor to automation, H_t^A, is low and automation intensity is low. Consequently, as stipulated in Proposition 3.B, growth is driven by horizontal innovation and the behavior of the economy is close to that of a Romer model with a Cobb-Douglas production function with low- and high-skill labor. Both wages approximately grow at a rate $g_t^N/(\sigma-1)$ and the labor share is approximately constant. G_t depreciates following equation (12).

Rising automation. As w_{Lt} grows relative to $\tilde{\varphi}^{-1}$, the term $(V_t^A - V_t^N)/(w_{Ht}/N_t)$ increases, which raises the incentive to innovate in automation. Without the externality in the automation technology ($\widetilde{\kappa} = 0$), (19) directly implies that H_t^A must rise significantly above zero, and with it the Poisson rate of automation, $\eta\left(H_t^A/(1-G_t)\right)^{\kappa}$ and thereby the share of automated products, G_t. For $\widetilde{\kappa} > 0$, the initial depreciation in

[24]Beyond a focus on different empirical phenomena (an increase in inequality vs. cross-country productivity differences), there are three important differences between our model and Zeira (1998). He assumes exogenous technical progress and focuses on adoption, while we model two types of endogenous innovation. The innovation cost changes over time in our model while the cost of adoption is 0 in Zeira (1998). As a result, automation is not labor-saving for the aggregate economy in his model.

the share of automated products gradually makes the automation technology less effective which delays the take-off of automation. Our initial assumption that knowledge spillovers are not too large ($\widetilde{\kappa} < 1-\kappa(1-\beta)$) is a sufficient (but not necessary) condition for the take-off to always happen.[25]

Following Proposition 1, the increases in G_t and N_t lead to an increase in the skill premium and a decline in the labor share.[26] For some parameters, low-skill wages temporarily decline (see numerical examples in Appendix 7.6.2).[27] Arguably, this is where our model differs the most from the rest of the literature, notably because a model with fixed G, a capital-skill complementarity model like KORV or a factor-augmenting technical change model with directed technical change as Acemoglu (1998) do not feature labor-saving innovation and therefore cannot lead to a decline in low-skill wages.

High but stable automation. With the share of automated products, G_t, no longer near zero, the gain from automation $V_t^A - V_t^N$ and its effective cost w_{Ht}/N_t grow at the same rate (the right-hand side in (21) is close to $1/G_t$). As a result, the normalized mass of high-skill workers in automation research ($N_t h_t^A$) stays bounded (see (19)), and so does the Poisson rate of automation, such that G_t converges to a positive constant below 1. The economy then converges toward an asymptotic steady-state where Proposition 2 applies. High-skill wages grow at the same rate as output and low-skill wages grow at a positive but lower rate while the labor share stabilizes again.

With positive growth in low-skill wages, the profits of non-automated firms become negligible relative to those of automated firms with $g_\infty^{\pi^N} = g_\infty^{\pi^A} - \beta(\sigma-1) g_\infty^{w_L}$. As total profits are proportional to output, the profits and value of automated firms grow at the rate of output minus the growth rate of the number of products: $g_\infty^{V^A} = g_\infty^{\pi^A} = g_\infty^Y - g_\infty^N$. Asymptotically, the value of a non-automated firm entirely lies in the possibility of future automation. Therefore, it grows at the same rate as the value of an automated firm: $g_\infty^{V^N} = g_\infty^{V^A}$. In other words, the prospect of future automation eventually guarantees the entry of new products.[28] The model continues then to differ from a generic capital

[25] Alternatively, if we assume that the automation technology obeys $max\left\{\eta G_t^{\widetilde{\kappa}}, \underline{\eta}\right\}\left(N_t h_t^A\right)^\kappa$ with $\underline{\eta} > 0$, then automation always take off.

[26] Changes in the mass of high-skill workers in production, H_t^P, also affect the skill premium and the labor share. Increasing G_t requires hiring more high-skill workers in automation innovation (in the same vein as the General Purpose Technology literature; notably Beaudry et al., 2016), but at the same time horizontal innovation declines on average during this time period, which increases H_t^P.

[27] Generating a decline in low-skill wages is harder with endogenous than exogenous technical change because a decline in low-skill wages reduces further automation, which is why the decline is temporary. Whether low-skill wages have declined in the data depends on how one accounts for compositional changes in the low-skill population, work benefits and the deflator (see Section 4.3).

[28] The economy would not admit such an asymptotic steady-state if automation was entirely under-

deepening model since the share of automated firms G_t is only constant thanks to the interplay between horizontal innovation and automation while the long-run growth rate depends on automation. Appendix 7.5.6 derives comparative statics in the steady-state and shows that the asymptotic growth rates of GDP and low-skill wages increase in the productivity of automation η and horizontal innovation γ.

Overall, our model predicts that we should see an increase in automation as an economy develops, consistent with the increase in automation innovations observed since the 1970s (as documented in Stylized fact 3). In line with the results of Section 2, this increase in automation is associated with an increase in the skill premium and a decline in the labor share, which have also been observed in the United States (Stylized facts 1 and 2). This contrasts our paper with most of the growth literature which relies on exogenous shocks to explains these trends. Section 4 shows that an extended version of our model can reproduce those trends not only qualitatively but also quantitatively.

3.4 Interactions between automation and horizontal innovation

Before moving to the quantitative exercise, we show how the interactions between automation and horizontal innovation can help account for two puzzles in the literature on inequality and technical change.

Increasing automation and decelerating skill premium. In recent years, the growth rate of the skill premium has declined (see Figure 1.A) while at the same time innovation has been more directed toward automation (Figure 1.C). At first glance, this seems to contradict an explanation of the increase in the skill premium by automation. Yet, in our model, there is no one-to-one link between the growth rate of the skill premium and the direction of innovation. First, the share of automated products G_t can be understood as a stock variable which increases with the automation of not-yet automated products but depreciates through horizontal innovation. As a result, maintaining a high level for G_t requires a high level of automation innovations. Typically, the skill premium rises the fastest when G_t increases sharply and decelerates when G_t approaches its steady-state value, which may, however, be when innovation is the most intensely directed toward automation (this happens both in Section 4 below and in the

taken by entrants replacing the incumbents. In that case, the value of creating a new product would only correspond to the discounted flow of profits of a non-automated firm, which grows slower than the cost of horizontal innovation (high-skill wages normalized by N_t) and horizontal innovation could not be sustained. In contrast, the steady-state would still exist if the incumbent also automated with positive probability or captured a share of the surplus created by an automation innovation.

numerical simulation of Appendix 7.6.1). Second, since our model does not feature a CES production with factor-augmenting technologies (as Katz and Murphy, 1992, Acemoglu, 1998, or Goldin and Katz, 2008), the elasticities of the skill premium with respect to the two technology variables (G_t and N_t) are not constant.

Automation with no increase in growth. A second puzzle is that as the skill-premium has increased, GDP growth has not accelerated, which casts doubt on whether a technological revolution is happening (see Acemoglu and Autor, 2011). Our model offers a potential explanation: horizontal innovation may decline when automation takes off, with an ambiguous net effect on growth. Formally, we find that the rate of horizontal innovation is lower in the steady-state than for low N_t (proof in Appendix 7.5.5):

Corollary 1. *For any G_0 there exists an N_0 sufficiently low, that the horizontal innovation rate, g^N is initially higher than in the asymptotic steady-state.*

Intuitively, three effects explain this result: First, once automation sets in, some high-skill workers are hired in automation research which reduces the amount of high-skill workers in production and therefore reduces horizontal innovation through a classic scale effect. Second, the elasticity of GDP growth with respect to horizontal innovation is larger in the asymptotic steady-state, which, from the Euler equation, increases the elasticity of the interest rate with respect to horizontal innovation and reduces horizontal innovation. Third, asymptotically, a new firm makes negligible profits relative to the cost of innovation until it gets automated, which further reduces horizontal innovation.

4 Quantitative Exercise

We conduct a quantitative exercise to compare empirical trends for the United States with the predictions of our model. We proceed in three steps. First, we calibrate our model, then we show that the data call for an increase in the share of automated products G and assess how well our model can do relative to a more flexible set-up with exogenous technology. Finally, we analyze the future evolution of the economy.

4.1 Extended model and calibration

To match the data quantitatively, we modify the baseline model. First, since the share of high-skill workers has dramatically increased, we let H and L vary over time and use the paths from the data. Second, we assume that producers rent machines from

a capital stock. Capital can also be used as structures in both automated and non-automated firms. Third, we allow for the possibility that low-skill workers are replaced by a composite of machines and high-skill workers. The production function (2) becomes:

$$y(i) = [l(i)^{\frac{\epsilon-1}{\epsilon}} + \alpha(i)(\tilde{\varphi} h_e(i)^{\beta_4} k_e(i)^{1-\beta_4})^{\frac{\epsilon-1}{\epsilon}}]^{\frac{\epsilon\beta_1}{\epsilon-1}} h_s(i)^{\beta_2} k_s(i)^{\beta_3}, \tag{22}$$

where $\beta_1 + \beta_2 + \beta_3 = 1$ and $\beta_4 \in [0, 1)$. The central difference between equations (2) and (22) is the introduction of $h_e(i)$ as high-skill labor which—along with machines—perform the newly automated tasks ("e" for equipment). This feature is necessary to capture a relatively low drop in the labor share. $k_s(i)$ is structures and $k_e(i)$ and $k_s(i)$ are both rented from the same capital stock K_t. K_t increases with investment in final goods and depreciates at a fixed rate Δ, so that:

$$\dot{K}_t = Y_t - C_t - \Delta K_t. \tag{23}$$

The cost advantage of automated firms now depends on the ratio between low-skill wage and the price of the high-skill labor capital aggregate, $w_{Lt}/(w_{Ht}^{\beta_4} \tilde{r}_t^{1-\beta_4})$, where $\tilde{r}_t = r_t + \Delta$ is the gross rental rate of capital. The logic of the baseline model directly extends to this case. Proposition 1 still holds and Proposition 2 holds with $g_\infty^{w_H} = g_\infty^Y = g_\infty^N / ((\beta_2 + \beta_1\beta_4)(\sigma - 1))$ and $g_\infty^{w_L} = g_\infty^Y (1 + (\sigma - 1)\beta_1\beta_4) / (1 + \beta_1(\sigma - 1))$. An equivalent to Proposition 3 holds but the system of differential equations includes three control variables and three state variables. The transitional dynamics are similar to that of the baseline model but automation innovation now depends on $w_{Lt}/(w_{Ht}^{\beta_4} \tilde{r}_t^{1-\beta_4})$: automation innovation is low when N_t is low, increases later on before stabilizing as the economy approaches its asymptotic steady-state. The capital share and the capital output ratio increase when automation increases as equipment replaces low-skill labor in production. Details and proofs are provided in Appendix 7.10.

We match our extended model to the data (see Appendix 7.11 for details). We identify low-skill workers with non-college educated workers and high-skill workers with college educated workers and focus on the years 1963-2012 (workers with "some college" are assigned 50/50 to each category following the methodology of Acemoglu and Autor, 2011). We match the skill-premium and take the empirical skill-ratio as given (and normalize total population to 1). We also match the growth rate of real GDP/employment and the labor share. We associate the use of machines with private equipment (excluding transport) and software. As pointed out by Gordon (1990) the NIPA price indices for

Table 1: Parameters from quantitative exercise

Parameter	σ	ϵ	β_1	γ	$\tilde{\kappa}$	θ	η	κ	ρ	β_2	Δ	β_4	$\tilde{\varphi}$	N_{1963}	G_{1963}
Value	5.94	4.24	0.59	0.44	0.56	1	0.40	0.65	0.034	0.18	0.011	0.73	1.57	9.9	0.03

real equipment are likely to understate quality improvements in equipment and therefore growth in the real stock of equipment. Hence, we use the adjusted price index from Cummins and Violante (2002) for equipment, and build (private) equipment and software to GDP ratios from 1963 to 2000.

Our model is not stochastic and cannot be directly estimated. Instead, we take a parsimonious approach and choose parameters to minimize the weighted squared log-difference between observed and predicted paths. We start the simulation 40 years before 1963 to force N_{1963} and G_{1963} to be consistent with the long-run behavior of our model.[29] Since the model requires the skill-ratio before and after the time period we estimate, we fit a generalized logistical function to the path of the log of the skill-ratio and use the predicted values outside 1963-2012 (over that time period, the fit is excellent).

The model features a total of 13 parameters with two initial conditions N_{1923} and G_{1923}. We allow all parameters to adjust freely (other than the economically motivated boundaries imposed by the model itself) and then assess whether these parameter estimates fit with other similar estimates. Table 1 gives the resulting parameters and the values of the state variables N_t and G_t in the first matched year, 1963. The elasticity of substitution across products σ is estimated at 5.94, in line with Christiano, Eichenbaum and Evans (2005) who find that observed markups are consistent with a value of around 6. The elasticity of substitution between machines and workers is estimated at 4.24. $\tilde{\kappa}$ is estimated at 0.56 implying a substantial automation externality; a force which accelerates the increase in the share of automated products. Finally, we find a β_1—the factor share of machines/low-skill workers—of 0.59 which implies sizable room for automation, though a β_4 of 0.73 means that the share of high-skill workers in the composite that replaces low-skill workers is of substantial importance. The preference parameters are within standard estimates with a ρ of 3.4% and the implied θ resulting in log-preferences. The only parameter that is estimated outside a common range is the depreciation rate Δ (though it is not precisely identified). Appendix 7.11.3 discusses in details how the parameters are identified.

[29]Initial values for G_t and K_t have little impact on the state of the economy several years later. By simulating the economy 40 years prior, we ensure that the simulated moments are nearly independent of the initial values for G_t and K_t. We fix K_{1923} at its steady-state value in a model with no automation.

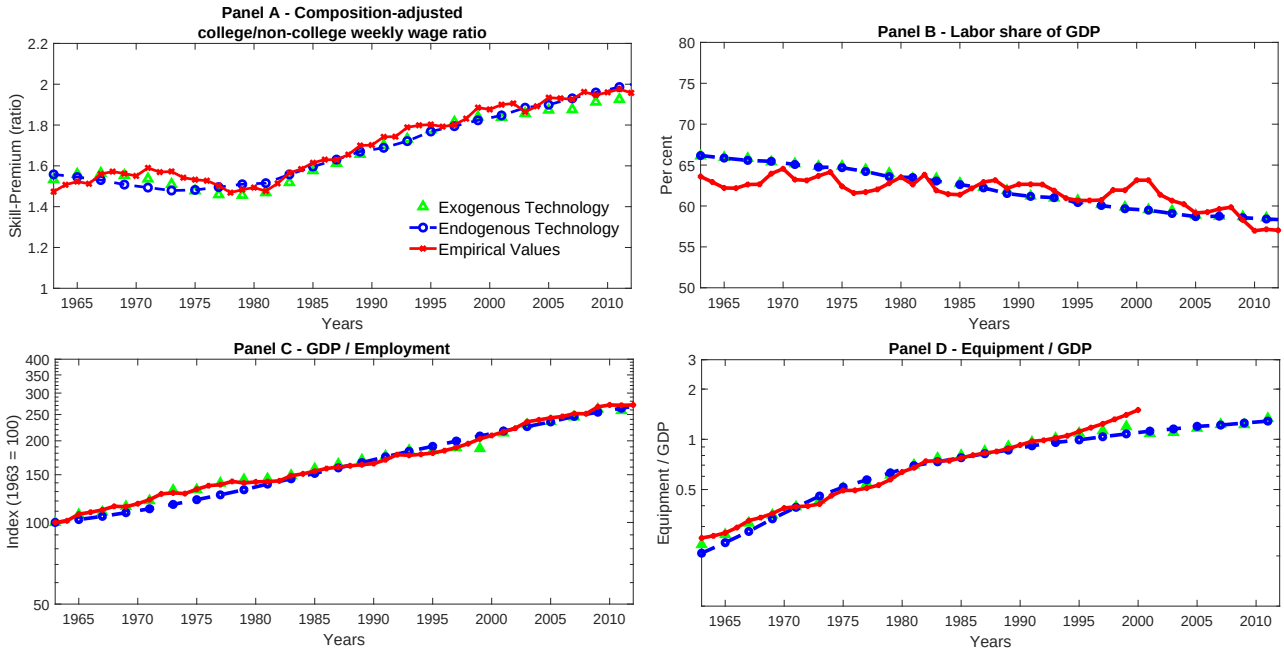

Figure 3: Empirical paths and predicted paths for the endogenous and exogenous technology models

Figure 3 further shows the predicted path of the matched data series ("endogenous technology" series—the "exogenous" series will be explained in Section 4.2) along with their empirical counterparts. Panel A demonstrates that the model matches the rise in the skill premium from the early 1980s and the flat skill premium in the period before reasonably well. Though a bit less pronounced than in the data, our model also includes the more recent decline in the growth rate of the skill premium, which, computed over a 5 years moving window, peaks in 1984 at 1.31% and drops to 0.52% in 2005. The total predicted decline in the labor share (8.7 p.p.) is slightly higher than the actual one (6.6 p.p.).[30] The average growth rate of the economy is matched completely as shown in Panel C. Although, the model largely captures the average growth rate of capital equipment over GDP during the period, the predicted path differs somewhat from its empirical counterpart as shown in Panel D on log-scale. Whereas the empirical path is close to exponential, the predicted path tapers off somewhat towards the end of the period.[31] Naturally, our model has a harder time capturing higher-frequency movements

[30]Our model sees the decline in the labor share as being a gradual phenomenon (instead of a sharp trend break from 2000) in line with the interpretation of Karabarbounis and Neiman (2013) and data on the global labor share.

[31]Recall that the ratio of equipment to GDP in the data is only a proxy for the ratio of machines to GDP in the model, since not all equipment is used to replace low-skill workers. Interestingly, more

such as a temporary increase in the labor share at the end of the 1990s. Further, to assess the predictive power of our model, we reproduce the same exercise but only matching the first 30 years. Appendix 6.4.1 reports the results: the parameters are nearly identical and the calibrated model performs well out-of-sample.

As shown in Table 1, the share of automated products G_t is low in 1963 at 2.5% far from its steady-state value of 92%. Figure 5 below shows that G_t increases sharply through the 1963-2012 time period and reaches 20% by 1986 and 64% by 2012. Therefore, our calibration strongly suggests that automation has risen in recent decades.

4.2 Alternative technological processes

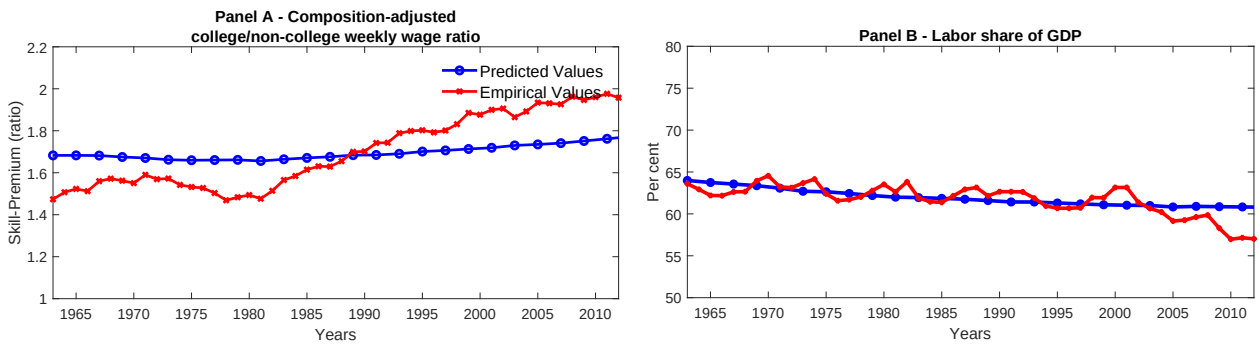

Figure 4: Empirical and predicted paths for a model with constant G.

Section 2.3 argued that a rise in the skill premium and a decrease in the labor share are consistent with an increase in the number of products together with either a rising or constant (but positive) share of automated products. It is therefore worth asking which features of the data require an increase in the share of automated products G_t. To answer this question, we consider an alternative model where G is a constant. We use the same procedure as above to calibrate this alternative economy but let G and the initial capital stock K_{1963} be free parameters. Figure 4 compares the empirical paths of the skill premium and the labor share with the ones predicted by a constant G model and shows that this alternative model does not match the data. Intuitively, a model with constant G has a hard time reconciling a roughly constant skill-premium followed by a fast rise without a sharp increase in economic growth. Appendix 7.11.4 shows the two other moments and reports the parameters.

recent data show a slow down in software investment (see Beaudry et al., 2016, Eden and Gaggl, 2018).

Our paper argues that the evolution of the income distribution in the United States can largely be accounted for with endogenous technical change. This is in contrast with an alternative view which stresses the importance of structural breaks in the data. To evaluate how well our endogenous innovation model performs, we compare it to a model with exogenous technology. We keep the (non-innovation) parameters from Table 1 but let N_t and G_t be free parameters between 1963 and 2012 chosen to match the empirical moments as close as possible.[32] Figure 3 shows the predicted paths for our moments in that case ("exogenous technology" series). This model performs slightly better than the endogenous one, notably when it comes to labor productivity fluctuations, but it does not capture trends better. Figure 5 compares the evolution of N_t and G_t in our endogenous growth model with this exogenous alternative. Panel B shows that the exogenous path for G_t also features a smooth but sharp increase from 1980 but the path of our endogenous growth model is very similar. Panel A shows that the exogenous path for N_t is more volatile than the endogenous path since the exogenous model tries to better capture the short-run fluctuations. The trends, however are similar. This pattern is robust to including exogenous labor-augmenting technical change (see Appendix 7.11.5).

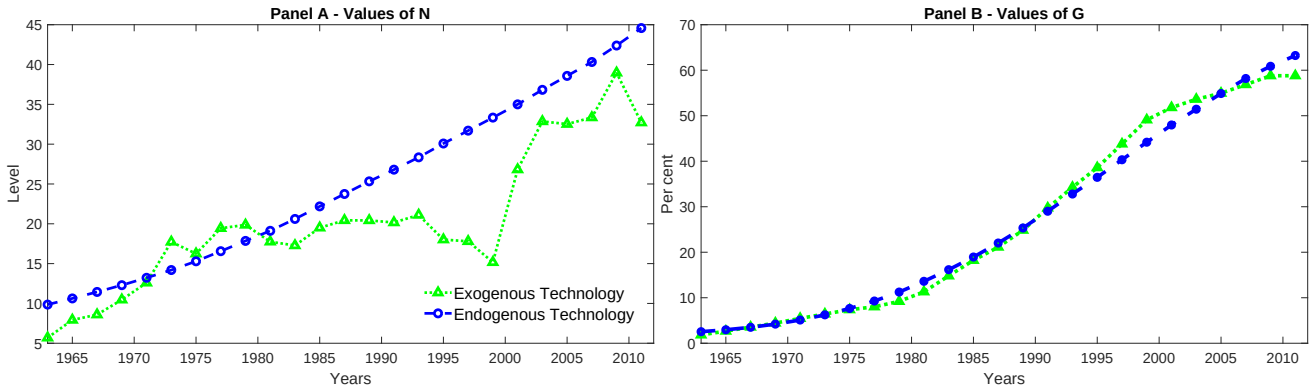

Figure 5: Paths for N_t and $G_t$in the endogenous and exogenous growth models.

Figure 5.B shows that the increase in G_t must be sharp to match the data, which is why our estimation procedure found a high value for the automation externality parameter $\tilde{\kappa} = 0.57$. Appendix 6.4.2 discusses this further by calibrating the model with $\tilde{\kappa} = 0$

[32]Since K_t still evolves endogenously, this corresponds to a Ramsey model with fully anticipated shocks to N_t and G_t. We let the initial capital stock K_{1963} be a free parameter as well. For the technology paths after 2012, we let N_t grow at a constant rate which is another free parameter and assume that G_t stays constant at its 2012 value. To facilitate comparison, we also assume that H_t^P in the exogenous growth model is fixed to be the same as in the endogenous growth model.

and showing that in this case, the endogenous and exogenous growth models deliver very different paths for G_t.

4.3 Evolution of the calibrated economy

Figure 6 further analyzes the behavior of the calibrated economy by plotting the transitional dynamics from 1963 to 2063. Panel A shows that GDP growth slows down past 2012 in line with recent economic trends—as argued in Section 3.4, our model can account for a slowdown in growth despite a high level of automation innovation thanks to a decline in horizontal innovation. Panels A and B show that the skill premium keeps growing albeit at a slower rate: over the 1980s the skill premium grew at an average of 1% a year according to the model and 0.7% in the 2000s, with a predicted growth rate of the skill premium of 0.2% in the 2050s as the share of automated products stabilizes. In the context of Proposition 1, for our parameters automation is always low-skill labor saving and horizontal innovation low-skill biased. In the meantime, the labor share smoothly declines toward its steady-state value of 52.8% and the high-skill labor share increases.[33] Panel C plots the share of automated products G_t, which keeps rising very slowly past 2012 toward its asymptotic value at 92%. The share of automation innovations is already above 40% in the mid1970s and increases steadily until the 2000s (Panel C). It peaks in 2005, even though the skill premium is decelerating. As argued in Section 3.4, this occurs in part because the level of automation can be thought of as a stock that depreciates with the entry of new products. This constitutes a response to the critique of the literature on SBTC put forward by Card and DiNardo (2002), who argue that inequality rising the most in the early to mid 1980s and technological change continuing in the 1990s, squares poorly with the predictions of a framework based on SBTC. The model predicts that the share of automation innovation will remain high in the future but at a slightly lower level than in the 2000s.

Acemoglu and Autor (2011) highlight that low-skill wages have declined since the 1980s, and we reproduce their series from 1963 to 2012 in Panel D. Our model does not distinguish between wages and other labor costs, so that any mention of wages so far should be understood as labor costs. At the same time, there has been a significant decline in the ratio of wages to total labor costs (from close to 1 to around 2/3 in

[33] With the decline in horizontal innovation, fewer high-skill workers are allocated to innovation activities overall, which also contributes to the decline in the labor share notably from the mid-1980s.

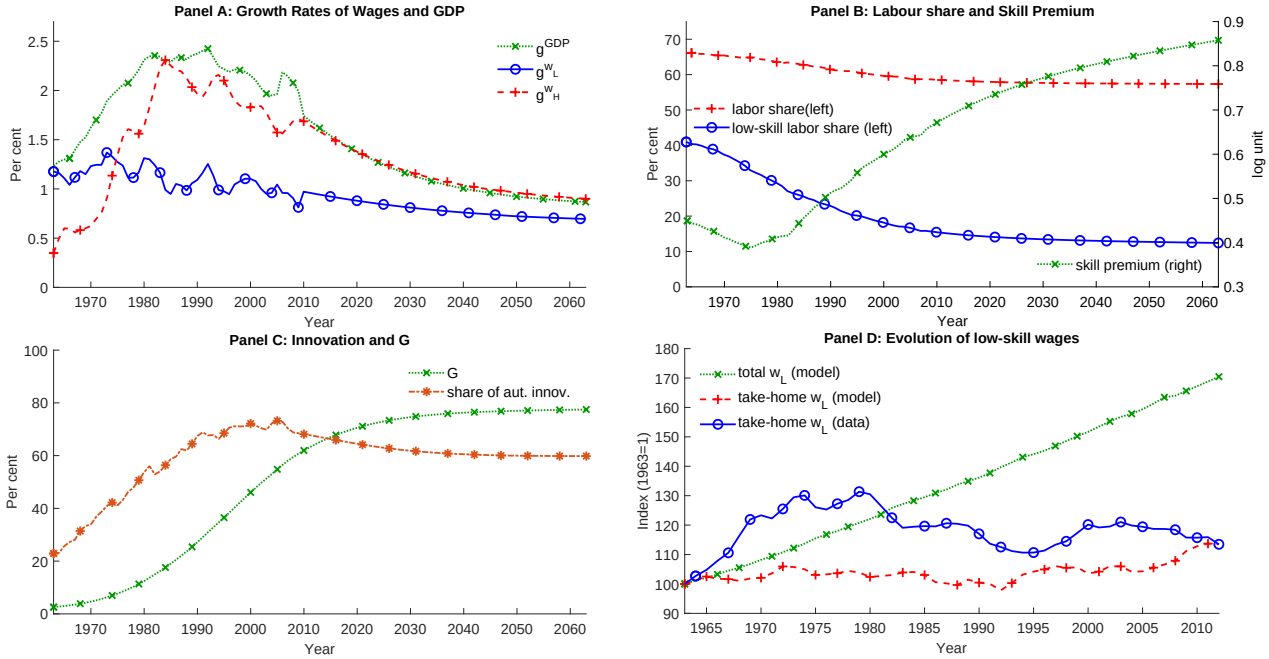

Figure 6: Transitional Dynamics with calibrated parameters (the growth rates are computed over a 5 year moving average).

our data).[34] Panel D plots our predicted low-skill labor costs (w_L) and the take-home low-skill wages (net of work benefits, social contributions, etc). Our model predicts a very slow growth for low-skill labor costs (70% over 50 years), which is consistent with stagnating take-home low-skill wages. Appendix 7.11.6 analyzes the effect of taxes on machines and automation innovations on low-skill wages and output.

Our exercise bears similarities to KORV who also seek to explain the increase in the skill premium using capital-skill complementarity while matching the labor share. There are four major differences. First, our exercise is more demanding since instead of directly feeding in the empirical path of equipment, we calibrate a general equilibrium model which endogenizes both capital and technology.[35] Second, KORV do not attempt to match the evolution of labor productivity: given the large increase in the stock of equipment capital, their model would have to feature a large simultaneous unexplained decrease in the growth rate of TFP. Third, their model does not match a decline in the labor share, but instead shows a slight increase toward the end of their sample period.

[34] Obtaining these figures requires combining several indices which are not perfectly consistent, nevertheless the trend is clear—see Appendix 7.11.2 for details.

[35] Others have tried to match similar trends, such as Eden and Gaggl (2018) using a model similar to KORV, or Goldin and Katz (2008) using the traditional model of SBTC with factor-augmenting technologies.

This is not an artifact of their calibration but a feature of their model. Their production function is a nested CES where low-skill labor is substitute with a CES aggregate of high-skill and equipment which are complement. Therefore, should the equipment stock keep rising (through investment-specific technological change), the income share of equipment would decline in the long-run.[36] Fourth, their model does not feature labor-saving innovation as an increase in investment specific technical change increases all wages for perfectly elastic capital (Appendix 7.12 elaborates on the last two points).

To summarize, our model makes the case that recent trends in the skill premium, the labor share, labor productivity and the ratio of equipment to GDP can be quantitatively explained as resulting from an endogenous increase in the share of automation innovations. This provides a different view of these trends as arising from the natural development of an economy instead of exogenous shocks. Furthermore, this viewpoint also explains why the rise in the skill premium was not accompanied by an increase in productivity growth and why the skill premium has decelerated despite a high share of automation innovations. Of course, our model cannot match the data perfectly. In particular, we do not capture high-frequency movements and we overestimate the decline of the labor share. Our exogenous N, G exercise shows that to better account for the data, one would have to modify the aggregate production function.

4.4 Data on automation innovations

We now provide some evidence based on patent data which suggests that the share of automation innovations has increased since the 1970s in line with our model. Classifying patents as automation versus non-automation is not straightforward and there are no technological codes in patent data aimed at doing so. Nevertheless, in a recent working paper, Mann and Püttmann (2018) classify US patents as automation versus non-automation using machine learning techniques (see Appendix 7.11.2 for details). Figure 7.A reports the shares of automation patents according to their analysis and according to our calibrated model. In line with our model, the share of automation patents according to their definition has markedly increased. Relative to their series, our model suggests a higher share of automation innovations (particularly at the beginning of the

[36] If low-skill labor were substitute with a Cobb-Douglas aggregate of high-skill and equipment (as in our automated firms), the capital share would increase in the long run. Yet, the growth rate of the skill premium would then increase when investment specific technological change accelerates, so such a model could not feature a deceleration in the skill premium when innovation seems to be the most biased against low-skill workers as in our model.

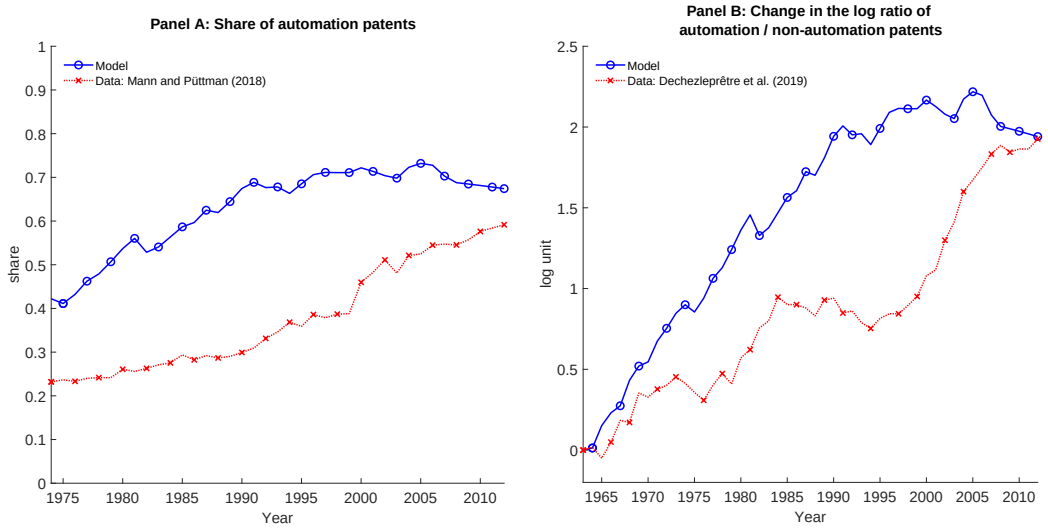

Figure 7: Trends in Automation innovations

sample), but the increase is of a similar magnitude.

Dechezleprêtre et al. (2019) offer an alternative classification of automation versus non-automation patents in machinery, which relies on the technological codes of patents and the presence of certain keywords in the text of patents (see Appendix 7.11.2 for details). They show that the share of automation patents in machinery is correlated with a decline in routine tasks across US industries and, using international firm-level data, that higher low-skill wages lead to more automation innovations but not more non-automation innovations. Their classification only allows for the identification of a subset of automation and non-automation patents. Yet, provided that these subsets are constant shares of both types of innovations, we can use the increase in the log ratio of their automation versus non-automation patents as a proxy for the increase of automation versus horizontal innovation in our model. We plot the model and data series (indexing the log ratios at 0 in 1963) in Figure 7.B. Here as well, we find similar trends—except for a small decline between 1984 and 1994 in the data. Any series based on patenting will be a noisy proxy for true underlying automation innovation. Despite this, both of these measures suggest that automation increased since the 70s and is still very high today, in line with the predictions of our model.

5 Conclusion

This paper introduces automation in a horizontal innovation growth model. We show that an increase in automation leads to an increase in the skill premium, a decline in the labor share and possibly a decline in low-skill wages. Moreover, such an increase in the share of automation innovations is the natural outcome of a growing economy since higher low-skill wages incentivize more automation innovations. Quantitatively, our model can replicate the evolution of the US economy since the 60s: a continuous increase in the skill premium with a more recent slowdown, a decline in the labor share, stagnating labor productivity growth and an increase in the share of automation innovations. We predict that the skill premium keeps rising in the future albeit at a lower rate and that the labor share stabilizes at a rate below today's.

The increase in the share of automation innovations which prompts changes in the income distribution occurs endogenously in our paper. This stands in contrast with 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 KORV, a change in the relative supply of skills, as per Acemoglu (1998), or the arrival of a general purpose technology as in the related literature. This feature is shared by 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.[37]

The present paper is only a step towards a better understanding of the links between automation, growth and income inequality. Given that automation has targeted either low- or middle-skill workers and that artificial intelligence may now lead to the automation of some high-skill tasks, a natural extension of our framework would include more skill heterogeneity. Another natural next step would be to add firm heterogeneity and embed our framework into a quantitative firm dynamics model. Our framework could also 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.

[37] This feature is also shared by Galor and Weil (2000) and Hansen and Prescott (2002) who endogenize the industrial revolution take-off.

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. (2002b). Directed Technical Change. *Review of Economic Studies*, 69(4):871–809.

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 Restrepo, P. (2017). Robots and jobs: Evidence from us labor markets.

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

Acemoglu, D. and Restrepo, P. (2018b). Demographics and automation. NBER working paper 24421.

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.

Aghion, P., Jones, B., and Jones, C. (2017). Artificial intelligence and economic growth.

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. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, 3.

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.

Beaudry, P. and Green, D. (2005). Changes in US wages, 1976-2000: ongoing skill bias or major technological change? *Journal of Labor Economics*, 23(3):609–648.

Beaudry, P., Green, D., and Sand, B. (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics*, 34(S1):S199–S247.

Benzell, S., Kotlikoff, L., LaGarda, G., and Sachs, J. (2017). Robots are us: some economics of human replacement. Working Paper WP-785.

Bloom, N., Schankerman, M., and van Reenen, J. (2013). Identifying technology spillovers and product market rivalary. *Econometrica*, 81(4):1347–1393.

Buera, F. and Kaboski, J. (2012). The Rise of the Service Economy. *American Economic Review*, 102(6):2540–2569.

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.

Christiano, E. and Evans (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1).

Clemens, M., Lewis, E., and Postel, H. (2018). Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion. *American Economic Review*, 108(6):1468–87.

Cummins, J. and Violante, G. (2002). Investment-specific technical change in the united states (1947-2000): Measurement and macroeconomic consequences. *Review of Economic Dynamics*, 5(2):243–284.

Dechezleprêtre, A., Hémous, D., Olsen, M., and Zanella, C. (2019). Automating labor: Evidence from firm-level patent data.

Eden, M. and Gaggl, P. (2018). On the welfare implications of automation. *Review of Economic Dynamics*, 29:15–43.

Feng, A. and Graetz, G. (2016). 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.

Galor, O. and Weil, D. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90(4):806–28.

Goldin, C. and Katz, L. (2008). *The Race between Education and Technology*. Harvard University Press.

Gordon, R. J. (1990). The measurement of durable goods prices. *NBER Monograph Series*.

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

Habakkuk, J. (1962). *American and British Technology in the Nineteenth Century*. Cambridge University Press.

Hansen, G. and Prescott, E. (2002). Malthus to solow. *American Economic Review*, 92(4):1205–1217.

Hémous, D. and Olsen, M. (2016). The rise of the machines: Automation, horizontal innovation and income inequality. CEPR Discussion Paper No 10244.

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.

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.

Mann, K. and Puttmann, L. (2018). Benign effects of automation: New evidence from patent texts.

Manuelli, R. and Seshadri, A. (2014). Frictionless technology diffusion: The case of tractors. *The American Economic Review*, 104(4):1368–1391.

Martinez, J. (2018). Automation, growth and factor shares.

Nakamura, H. and Zeira, J. (2018). Automation and unemployment: Help is on the way.

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.

Rahman, A. (2017). Rise of the machines redux: Education, technological transition and long-run growth.

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.

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.

6 Main Appendix

6.1 Proof of Proposition 1

We focus on the imperfect substitute case and Appendix 7.1.1 deals with the perfect substitute case. Rewrite (7) as

$$\frac{w_H}{w_L} = \frac{1-\beta}{\beta}\frac{L}{H^P}\frac{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+1-G}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G}. \tag{24}$$

Since $0<\mu<1$, (24) establishes w_H as a function of G, H^P and w_L such that w_H is increasing in w_L and G and decreasing in H^P, with $w_H/w_L > (1-\beta)/\beta \times L/H^P$ for $G>0$. (6) similarly establishes w_H as a function of N, G and w_L which is decreasing in w_L and increasing in N and G. w_H, w_L are then jointly uniquely determined by (24) and (6) for given N, G and H^P. Both increase in N, and w_H increases in G. In addition (6) traces a convex iso-cost curve in the input prices plan.

In addition, (24) shows that w_H/w_L increases with w_L. Since w_L increases in N, then w_H/w_L increases in N as well. If w_L decreases in G, then since w_H increases in G, w_H/w_L increases in G. If instead w_L increases in G, then the right-hand side of (24) increases with G both directly and because w_L increases, this ensures that w_H/w_L still increases in G. Therefore both an increase in N and an increase in G are skill-biased and following (9) decrease the labor share. This establishes Parts A and B except for the relationship between w_L and G.

Comparative statics of w_L with respect to G. Combine (24) and (6) to get:

$$w_L = \begin{array}{c} \frac{\sigma-1}{\sigma}\beta\left(\frac{H^P}{L}\right)^{(1-\beta)} N^{\frac{1}{\sigma-1}}\left(G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+(1-G)\right)^{1-\beta} \\ \times\left(G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)\right)^{\frac{1}{\sigma-1}-(1-\beta)} \end{array}, \tag{25}$$

Log differentiating with respect to G one obtains:

$$\widehat{w}_L = \left[\begin{array}{c} \underbrace{\frac{1}{\sigma-1}\frac{\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)}}_{\text{scale effect}} \\ \underbrace{-(1-\beta)\left(\frac{1-\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+(1-G)}+\frac{\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)}\right)}_{\text{substitution effect}} \end{array}\right]\frac{G\widehat{G}}{Den} \tag{26}$$

$$\text{where } Den \equiv \begin{array}{c} 1-\frac{\beta\varphi w_L^{\epsilon-1}}{\left(1+\varphi w_L^{\epsilon-1}\right)}\frac{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)} \\ +\frac{\varphi w_L^{\epsilon-1}(\epsilon-1)(1-\beta)}{\left(1+\varphi w_L^{\epsilon-1}\right)}\left(\frac{\mu G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+1-G}+\frac{(1-\mu)G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G}\right) \end{array}.$$

$Den > 0$ as $\epsilon > 1$, $\mu \in (0,1)$ and $\frac{\beta\varphi w_L^{\epsilon-1}}{1+\varphi w_L^{\epsilon-1}}\frac{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)} < 1$. In (26) the scale effect term is positive as $\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1>0$. This term comes from the differentiation of (6) with respect to G at constant w_H (hence it represents the shift right of the isocost curve). The substitution effect term is negative because $1-\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}>0$ since $\mu<1$, it comes from the differentiation of (7) with respect to G.

First note that if $\frac{1}{\sigma-1} \leq 1-\beta$, the scale effect is always dominated by the substitution effect. Hence w_L is decreasing in G.

If instead $\frac{1}{\sigma-1} > 1-\beta$, the scale effect is dominated by the substitution effect provided that $R_W \equiv \frac{1-\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G} / \frac{\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+1-G}$ is large enough. From (25) we get:

$$w_L = \frac{\sigma-1}{\sigma}\beta N^{\frac{1}{\sigma-1}}\left(\frac{H^P}{L}\frac{1-G\left(1-\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}\right)}{G\left(\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1\right)+1}\right)^{1-\beta}\left(G\left(\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1\right)+1\right)^{\frac{1}{\sigma-1}}.$$

Using that $G \in [0,1]$, we obtain that

$$w_L\left(1+\varphi w_L^{\epsilon-1}\right)^{1-\beta} > \frac{\sigma-1}{\sigma}\beta\left(H^P/L\right)^{1-\beta}N^{\frac{1}{\sigma-1}},$$

which ensures that $\lim_{N\to\infty} w_L = \infty$ uniformly with respect to G (i.e. for any $\overline{w}_L > 0$, there exist $\overline{N}$ such that for any $N > \overline{N}$ and any G, $w > \overline{w}_L$). Since $\lim_{w_L\to\infty, G\to 1} R_W = \infty$, we get that $\lim_{N\to\infty, G\to 1} R_W = \infty$, so that for N and G large enough the substitution effect dominates. This establishes Part B.

Part C. Log-differentiating (25) with respect to N gives $\widehat{w}_L = \frac{\widehat{N}}{\sigma-1}\frac{1}{Den}$. We then combine this expressions with (26) and use that for an increase in the number of non-automated products only, NG is a constant so that $\widehat{G} = -\widehat{N}$. Denoting $\widehat{w}_L^{NT}$ (NT for "new tasks"), the change in w_L, we then get:

$$\widehat{w}_L^{NT} = \left[\begin{array}{c} \frac{(1-\beta)G\left(\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1\right)}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)}+\frac{(1-\beta)G\left(1-\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}\right)}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+(1-G)} \\ +\frac{(\sigma-1)^{-1}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)} \end{array}\right]\frac{\widehat{N}}{Den}. \tag{27}$$

Hence low-skill wages always increase with the arrival of non-automated products. Log-differentiating (7), one gets:

$$\widehat{w}_H - \widehat{w}_L = \begin{array}{c} \left(\frac{\mu G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+1-G} + \frac{(1-\mu)G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G} \right) \frac{(\epsilon-1)\varphi w_L^{\epsilon-1}}{1+\varphi w_L^{\epsilon-1}} \widehat{w}_L \\ + \left(\frac{G\left(\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1\right)}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+1-G} + \frac{G\left(1-\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}\right)}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G} \right) \widehat{G} \end{array}$$

Using (27) and that $\widehat{G} = -\widehat{N}$, we get that following an increase in the mass of non-automated products (keeping NG constant):

$$\begin{aligned} & \widehat{w}_H^{NT} - \widehat{w}_L^{NT} \\ = & \left[\begin{array}{c} \left(\frac{\mu\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+1-G} + \frac{(1-\mu)\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G} \right) \frac{(\epsilon-1)\varphi w_L^{\epsilon-1}}{1+\varphi w_L^{\epsilon-1}} \frac{(\sigma-1)^{-1}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)} \\ - \left(\frac{\left(\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}-1\right)}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+1-G} + \frac{\left(1-\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}\right)}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G} \right) \left(1 - \frac{\beta\varphi w_L^{\epsilon-1}}{\left(1+\varphi w_L^{\epsilon-1}\right)} \frac{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)} \right) \end{array} \right] \frac{G\widehat{N}}{Den} \end{aligned}$$

which simplifies into:

$$\widehat{w}_H^{NT} - \widehat{w}_L^{NT} = \frac{\left[\left(\frac{\epsilon-1}{\sigma-1} - \beta \right) \frac{1}{1+\varphi w_L^{\epsilon-1}} - (1-\beta) \right] \left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1} \varphi w_L^{\epsilon-1}}{\left(G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+1-G\right)\left(G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G\right)} \frac{G\widehat{N}}{Den}. \tag{28}$$

Therefore an increase in the mass of non-automated products reduces the skill premium (and increases the labor share) if and only if $1-\beta > \left(\frac{\epsilon-1}{\sigma-1} - \beta\right) / \left(1+\varphi w_L^{\epsilon-1}\right)$. This in turn is true for w_L sufficiently large (that is N large enough) or for $\epsilon < \sigma$.

Combining (28) with (27), we further get

$$\widehat{w}_H^{NT} = \frac{(\sigma-1)^{-1}}{G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu}+(1-G)} \left(1 + \frac{G(1-\mu)\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}}{\left(G\left(1+\varphi w_L^{\epsilon-1}\right)^{\mu-1}+1-G\right)} \frac{(\epsilon-1)\varphi w_L^{\epsilon-1}}{1+\varphi w_L^{\epsilon-1}} \right) \frac{\widehat{N}}{Den}.$$

Therefore an increase in the mass of non-automated products leads to higher high-skill wages. This establishes Part C.

6.2 Proof of Proposition 2

To see that w_{Lt} is bounded from below, assume that liminf w_{Lt}=0. Then as $H_t^P$and G_t admit positive limits, (7) implies that liminf $w_{Ht} = 0$. Plugging this in (6) gives

liminf $N_t = 0$, which is impossible since g_t^N admits a positive limit. Therefore, w_{Lt} must be bounded below, so that (6) gives $g_\infty^{w_H} = \psi g_\infty^N$ where $\psi \equiv ((1-\beta)(\sigma-1))^{-1}$. Further, using that H_t^P admits a limit and that

$$w_{Ht} H_t^P = (1-\beta)\frac{\sigma-1}{\sigma} Y_t. \tag{29}$$

gives the growth rate of Y_t. To derive the asymptotic growth rate of w_{Lt} we consider in turn the cases $\epsilon < \infty$ and $\epsilon = \infty$.

Subcase with $\epsilon < \infty$. We use equation (25) which gives w_{Lt} as a function of N_t, G_t and H_t^P. Assuming that liminf w_{Lt} is finite leads to a contradiction, so that w_{Lt} tends to ∞. Then (25) implies that for $G_\infty < 1$,

$$w_{Lt} \sim \left(\left(\frac{\sigma-1}{\sigma}\beta \right)^{\frac{1}{1-\beta}} (1-G_\infty) \frac{H_\infty^P}{L} \left(G_\infty \varphi^\mu\right)^{\psi-1} \right)^{\frac{1}{1+\beta(\sigma-1)}} N_t^{\frac{\psi}{1+\beta(\sigma-1)}},$$

where $x_t \sim y_t$ signifies $x_t/y_t \to 1$. This delivers (11).

Subcase with $\epsilon = \infty$. Low skill wages are now defined as described in Appendix 7.1.1, and equation (11) immediately follows from $G_\infty < 1$.

6.3 Analytical Appendix to Section 3

In this Appendix we derive the system of normalized equation which characterizes the equilibrium and prove Proposition 6.3. Appendix 7.5 contains the proofs of intermediate results, of Corollary 1 and additional results mentioned in the text.

6.3.1 Normalized system of differential equations

Following Proposition 2, high-skill wages, output and consumption grow asymptotically proportionately to N_t^ψ with $\psi = ((1-\beta)(\sigma-1))^{-1}$ when $g_\infty^N > 0$ and $G_\infty > 0$. Therefore to study the behavior of the system, we introduce the normalized variables $\hat{v}_t \equiv w_{Ht} N_t^{-\psi}$ and $\hat{c}_t \equiv c_t N_t^{-\psi}$. As h_t^A mechanically tends to 0 as the mass of non-automated firms grows, we introduce $\hat{h}_t^A \equiv N_t h_t^A$. We define $\chi_t \equiv \hat{c}_t^{\,\theta}/\hat{v}_t$ which allows to simplify the system (χ_t is related to the mass of high-skill workers in production and therefore, given $\hat{h}_t^A$, to H_t^D and g_t^N). Since the economy does not feature a non-asymptotic steady-state, we need to keep track of the level of N_t by introducing $n_t \equiv N_t^{-\beta/[(1-\beta)(1+\beta(\sigma-1))]}$, which tends toward 0 as N_t tends toward infinity. Finally, we

define $\omega_t \equiv (w_{Lt} N_t^{-\psi/(1+\beta(\sigma-1))})^{\beta(1-\sigma)}$ which asymptotes a finite positive number.

We now derive the system of differential equations satisfied by the normalized variables (n_t, G_t, h_t, χ_t). By definition we get:

$$\dot{n}_t = -\frac{\beta}{(1-\beta)(1+\beta(\sigma-1))} g_t^N n_t. \tag{30}$$

Rewriting (12) with $\hat{h}_t^A$ gives:

$$\dot{G}_t = \eta G_t^{\widetilde{\kappa}} \left(\hat{h}_t^A\right)^{\kappa} (1-G_t) - G_t g_t^N. \tag{31}$$

Defining normalized profits $\widehat{\pi}_t^A \equiv N_t^{1-\psi} \pi_t^A$ and $\widehat{\pi}_t^N \equiv N_t^{1-\psi} \pi_t^N$ and the normalized values of firms $\widehat{V}_t^A \equiv N_t^{1-\psi} V_t^A$ and $\widehat{V}_t^N \equiv N_t^{1-\psi} V_t^N$, we can rewrite (14) and (15) as

$$\left(r_t - (\psi-1)\, g_t^N\right) \widehat{V}_t^A = \widehat{\pi}_t^A + \dot{\widehat{V}}_t^A, \tag{32}$$

$$\left(r_t - (\psi-1)\, g_t^N\right) \widehat{V}_t^N = \widehat{\pi}_t^N + \eta G_t^{\widetilde{\kappa}} \left(\widehat{h}_t^A\right)^{\kappa} \left(\widehat{V}_t^A - \widehat{V}_t^N\right) - \widehat{v}_t \widehat{h}_t^A + \dot{\widehat{V}}_t^N. \tag{33}$$

Equation (16) can similarly be rewritten as:

$$\kappa \eta G_t^{\widetilde{\kappa}} \left(\widehat{h}_t^A\right)^{\kappa-1} \left(\widehat{V}_t^A - \widehat{V}_t^N\right) = \widehat{v}_t. \tag{34}$$

Equation (17) implies that $\widehat{V}_t^N = \widehat{v}_t/\gamma$, therefore using (34) into (33), we get:

$$\left(r_t - (\psi-1)\, g_t^N\right) \widehat{v}_t = \gamma \widehat{\pi}_t^N + \gamma \frac{1-\kappa}{\kappa} \widehat{v}_t \widehat{h}_t^A + \dot{\widehat{v}}_t. \tag{35}$$

Since $w_{Lt}^{\beta(1-\sigma)} = \omega_t n_t$, we have using (8)

$$\widehat{\pi}_t^N = \omega_t n_t \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu} \widehat{\pi}_t^A. \tag{36}$$

Combining (31)-(36), we derive in Appendix 7.5.1:

$$\begin{aligned} \dot{\widehat{h}}_t^A = & \frac{\gamma \widehat{h}_t^A \omega_t n_t \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu} \widehat{\pi}_t^A}{(1-\kappa)\widehat{v}_t} + \frac{\gamma \left(\widehat{h}_t^A\right)^2}{\kappa} - \frac{\kappa \eta G_t^{\widetilde{\kappa}} \left(\widehat{h}_t^A\right)^{\kappa} \left(1 - \omega_t n_t \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu}\right) \widehat{\pi}_t^A}{(1-\kappa)\widehat{v}_t} \\ & + \eta G_t^{\widetilde{\kappa}} \left(\widehat{h}_t^A\right)^{\kappa+1} + \frac{\widetilde{\kappa} \widehat{h}_t^A}{1-\kappa} \left(\eta G_t^{\widetilde{\kappa}-1} \left(\hat{h}_t^A\right)^{\kappa} (1-G_t) - g_t^N\right). \end{aligned} \tag{37}$$

Rewriting (18) leads to $r_t = \rho + \theta \dot{\widehat{c}}_t / \widehat{c}_t + \theta \psi g_t^N$. Using this with (35) and (36), we get:

$$\dot{\chi}_t = \chi_t \left(\gamma \omega_t n_t \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{-\mu} \frac{\widehat{\pi}_t^A}{\hat{v}_t} + \gamma \frac{1-\kappa}{\kappa} \widehat{h}_t^A - \rho - (\theta\psi - \psi + 1) g_t^N \right). \tag{38}$$

Together equations (30), (31), (37) and (38) form a system of differential equations which depends on ω_t, $\widehat{\pi}_t^A / \widehat{v}_t$ and g_t^N. To determine $\widehat{\pi}_t^A / \widehat{v}_t$, note that profits are given by

$$\pi \left(w_L, w_H, \alpha \left(i \right) \right) = \frac{(\sigma - 1)^{\sigma - 1}}{\sigma^\sigma} c \left(w_L, w_H, \alpha \left(i \right) \right)^{1-\sigma} Y. \tag{39}$$

Using (3) and the definition of ω_t, one gets:

$$\pi_t^A = \frac{(\sigma - 1)^{\sigma - 1}}{\sigma^\sigma} \left(\beta^\beta (1-\beta)^{1-\beta} \right)^{\sigma - 1} \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^\mu w_{Ht}^{-\psi^{-1}} Y_t. \tag{40}$$

Rearranging terms in (6) gives

$$\widehat{v}_t = \left(\frac{\sigma - 1}{\sigma} \right)^{\frac{1}{1-\beta}} \beta^{\frac{\beta}{1-\beta}} (1-\beta) \left(G \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^\mu + (1-G) \omega_t n_t \right)^\psi. \tag{41}$$

Using the relationship between the high-skill wage bill and output in (29), we get:

$$Y_t = \sigma \psi \widehat{v}_t H_t^P N_t^\psi. \tag{42}$$

Therefore, rewriting (40) with (41) and (42), one gets:

$$\frac{\widehat{\pi}_t^A}{\widehat{v}_t} = \frac{\psi \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^\mu H_t^P}{G \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^\mu + (1-G) \omega_t n_t}, \tag{43}$$

which still requires finding H_t^P. Using (3), (4), $x/l = \varphi w_L^\epsilon$ and aggregating over all automated firms, one obtains the total demand of machines:

$$X_t = \beta G_t N_t \varphi \left(\frac{\sigma - 1}{\sigma} \right)^\sigma \left(\beta^\beta (1-\beta)^{(1-\beta)} \right)^{\sigma - 1} \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu - 1} w_{Ht}^{-\psi^{-1}} Y_t.$$

Using (41), this expression can be rewritten as:

$$X_t = \frac{\sigma - 1}{\sigma} \frac{\beta G_t \varphi \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu-1}}{G\left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + (1-G)\,\omega_t n_t} Y_t. \tag{44}$$

This together with (42) implies that $\widehat{c}_t$ obeys

$$\widehat{c}_t = \left(1 - \frac{\sigma - 1}{\sigma} \frac{\beta G_t \varphi \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu-1}}{G\left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + (1-G)\,\omega_t n_t}\right) \sigma \psi \widehat{v}_t H_t^P.$$

Combining this equation with (41), leads to

$$H_t^P = \frac{\left(\beta^{\beta} \frac{\sigma-1}{\sigma}\right)^{\frac{\theta^{-1}-\beta}{1-\beta}} (1-\beta)^{\frac{1}{\theta}} \chi_t^{\frac{1}{\theta}} \left(G_t \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + (1-G_t)\omega_t n_t\right)^{\psi\left(\frac{1}{\theta}-1\right)+1}}{G_t\left(\left(1-\beta\frac{\sigma-1}{\sigma}\right)\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)\left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu-1} + (1-G_t)\omega_t}. \tag{45}$$

Using the definition of H_t^D, one can rewrite (13) for high-skill workers as:

$$g_t^N = \gamma \left(H - H_t^P - (1-G_t)\,\widehat{h}_t^A\right). \tag{46}$$

Together (43), (45) and (46) determine $\widehat{\pi}_t^A / \widehat{v}_t$ and g_t^N as a function of the original variables $n_t, G_t, \widehat{h}_t^A, \chi_t$ and of ω_t, which still needs to be determined. To do so, combine (7) and (6) to obtain an implicit definition of ω_t:

$$\omega_t = \left[\begin{array}{c} \left(\frac{\sigma-1}{\sigma}\beta\right)^{\frac{1}{1-\beta}} \frac{H_t^P}{L} \left(G_t \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu-1} (\omega_t n_t)^{\frac{1-\mu}{\mu}} + (1-G_t)\right) \\ \times \left(G_t \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + (1-G_t)\omega_t n_t\right)^{\psi-1} \end{array}\right]^{\frac{\beta(1-\sigma)}{1+\beta(\sigma-1)}}. \tag{47}$$

Therefore, the system of differential equations satisfied by $n_t, G_t, \widehat{h}_t^A, \chi_t$ is defined by (30), (31), (37) and (38), with $\widehat{\pi}_t^A / \widehat{v}_t$, H_t^P, g_t^N and ω_t given by (43), (45), (46) and (47). The state variables are n_t and G_t and the control variables $\hat{h}_t^A$ and χ_t.

6.3.2 Proof of Proposition 3.A

To prove Proposition 3.A, we show that the system describes in Section 6.3.1 admits a steady-state:

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

$$\kappa^{-\kappa} \left(\gamma(1-\kappa)/\rho\right)^{\kappa-1} \rho/\eta + \rho/\gamma < \psi H. \tag{48}$$

Proof. We look for a steady state with positive long-run growth ($\left(g^N\right)^* > 0$) for the system defined by (30), (31), (37) and (38) and denote such a (potential) steady state $n^*, G^*, \widehat{h}^{A*}, \chi^*$ (we denote all variables at steady state with a *). Following (30), we immediately get $n^* = 0$. Using (31), G^* obeys:

$$G^* = \frac{\eta \left(G^*\right)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa}}{\eta \left(G^*\right)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + g^{N*}}. \tag{49}$$

We focus on a solution with $G^* > 0$ (when $\widetilde{\kappa} > 0$, $G^* = 0$ is also a solution), (49) implies that with $\left(g^N\right)^* > 0$, $G^* < 1$. Then, with $\mu \in (0,1)$, (47), implies that:

$$\omega^* = \left[\left(\tfrac{\sigma-1}{\sigma}\beta\right)^{\frac{1}{1-\beta}} \tfrac{H^{P*}}{L} \left(1 - G^*\right) \left(G^* \varphi^{\mu}\right)^{\psi-1} \right]^{\frac{\beta(1-\sigma)}{1+\beta(\sigma-1)}}.$$

Using (43), (38) implies that in steady state,

$$\widehat{h}^{A*} = \frac{\kappa}{\gamma\left(1-\kappa\right)} \left(\rho + \left(\left(\theta - 1\right)\psi + 1\right) g^{N*}\right) \tag{50}$$

which uniquely defines defines $\widehat{h}^{A*}$ as an increasing function of g^{N*} (recall that $\theta \geq 1$) with $\widehat{h}^{A*} > 0$ if $g^{N*} > 0$. Then, for $G^* > 0$, (49) combined with (50), defines G^* uniquely as an increasing function of g^{N*}. (46) also uniquely defines H^{P*} as a function of g^{N*}:

$$H^{P*} = H - \frac{g^{N*}}{\gamma} - \left(1 - G^*\right) \widehat{h}^{A*}. \tag{51}$$

(43) and (49) allows to rewrite (37) in steady state as:

$$\frac{\eta\kappa \left(G^*\right)^{\widetilde{\kappa}-1} \left(\widehat{h}^{A*}\right)^{\kappa}}{1-\kappa} \psi H^{P*} = \frac{\gamma}{\kappa} \left(\widehat{h}^{A*}\right)^2 + \eta G_t^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa+1}. \tag{52}$$

Since G^*, $\widehat{h}^{A*}$ and H^{P*} are functions of g^{N*}, one can rewrite (52) as an equation deter-

mining g^{N*}. A steady state with positive growth-rate is a solution to

$$f\left(g^{N*}\right) \equiv \frac{1-\kappa}{\kappa}\frac{\gamma G^{*}\widehat{h}^{A*}}{\psi H^{P*}}\left(\frac{1}{\kappa\eta\left(G^{*}\right)^{\widetilde{\kappa}}}\left(\widehat{h}^{A*}\right)^{1-\kappa}+\frac{1}{\gamma}\right)=1, \tag{53}$$

with $g^{N*}>0$. Indeed, (45) simply determines χ^{*} as:

$$\chi^{*}=\left(\frac{\sigma}{\sigma-1}\right)^{\frac{1}{1-\beta}(1-\theta\beta)}\frac{\left(1-\beta\frac{\sigma-1}{\sigma}\right)^{\theta}\left(H^{P*}\right)^{\theta}}{\left(1-\beta\right)\beta^{\frac{\beta}{1-\beta}(1-\theta)}\left(G^{*}\varphi^{\mu}\right)^{\psi(1-\theta)}}, \tag{54}$$

which achieves the characterization of a steady state for the system of differential equations defined by (30), (31), (37) and (38).

To establish the sufficiency of equation (48) for positive growth, note that as $g^{N*}\to 0$, then equations (50), (49) and (51) imply that

$$f\left(0\right)=\frac{\rho}{\psi H}\left(\frac{1}{\eta\kappa^{\kappa}\left(1-\kappa\right)^{1-\kappa}}\left(\frac{\rho}{\gamma}\right)^{1-\kappa}+\frac{1}{\gamma}\right).$$

In addition, $\frac{g^{N*}}{\gamma}+(1-G^{*})\,\widehat{h}^{A*}$ is always greater than $\frac{g^{N*}}{\gamma}$, therefore for a sufficiently large g^{N*} (smaller than γH), H^{P*} is arbitrarily small, while for the same value G^{*} and $\widehat{h}^{A*}$ are bounded below and above. This establishes that for g^{N*} large enough, $f\left(g^{N*}\right)>1$. Therefore a sufficient condition for the existence of at least one steady state with positive growth and positive G^{*} is that $f\left(0\right)<1$ (such that $f\left(g^{N*}\right)=1$ has a solution), which is equivalent to condition (48). The assumption that $\theta\geq 1$ further ensures that the transversality condition always holds. □

The steady state $(n^{*},G^{*},\hat{h}^{A*},\chi^{*})$ corresponds to an asymptotic steady state for our original system of differential equations (as $N_t\to\infty$ when $n_t\to 0$). Since $G_{\infty}=G^{*}\in(0,1)$, $g^{N}_{\infty}=g^{N*}>0$ and $H^{P}_{\infty}=H^{P*}>0$, Proposition 2 applies (which is also in line with the normalized variables being constant in steady-state). Section 7.5.4 provides further details on the behavior of the economy close to the steady-state.

6.3.3 Proof of Proposition 3.B

Here we prove Proposition 3.B. Combining (20) and (19), we can write:

$$N_t h^{A}_{t}=\left(\kappa\eta G_t^{\widetilde{\kappa}}\left(\int_t^{\infty}\exp\left(-\int_t^{\tau}r_u du\right)\left(\frac{N_t}{w_{Ht}}\left(\pi^{A}_{\tau}-\pi^{N}_{\tau}\right)d\tau-\frac{1-\kappa}{\kappa}\frac{N_t}{N_\tau}\frac{w_{H\tau}}{w_{Ht}}\left(N_\tau h^{A}_{\tau}\right)\right)d\tau\right)\right)^{\frac{1}{1-\kappa}}.$$

Using (29) and that aggregate profits $\Pi_t = N_t \left(G_t \pi_t^A + (1-G_t)\pi_t^N\right)$ are a share $1/\sigma$ of output, we can rewrite this equation as:

$$\widehat{h}_t^A = \left(\kappa\eta G_t^{\widetilde{\kappa}} \left(\int_t^\infty \exp\left(-\int_t^\tau r_u du\right)\left(\psi H_t^P \frac{\pi_\tau^A - \pi_\tau^N}{G_t\pi_t^A + (1-G_t)\pi_t^N} d\tau - \frac{1-\kappa}{\kappa}\frac{N_t}{N_\tau}\frac{\hat{v}_\tau}{\hat{v}_t}\widehat{h}_\tau^A\right) d\tau\right)\right)^{\frac{1}{1-\kappa}}. \tag{55}$$

Recalling (8), we can write:

$$\widehat{h}_t^A = \left(\kappa\eta G_t^{\widetilde{\kappa}} \left(\int_t^\infty \left(\begin{array}{c} \psi H_t^P \frac{\left(1+\varphi w_{L\tau}^{\epsilon-1}\right)^\mu - 1}{G_t\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^\mu + 1} \exp\left(\int_t^\tau \left(g_u^{\pi^N} - r_u\right) du\right) \\ -\frac{1-\kappa}{\kappa}\exp\left(\int_t^\tau \left(g_u^v - g_u^N - r_u\right) du\right)\widehat{h}_\tau^A \end{array}\right) d\tau\right)\right)^{\frac{1}{1-\kappa}}.$$

Consider a fixed $\widehat{t} > 0$. For an arbitrarily large T, if w_{L0} is sufficiently small relative to $\widetilde{\varphi}^{-1}$, w_{Lt} remains small relative to $\widetilde{\varphi}^{-1}$ over $\left(0, \widehat{t}+T\right)$. For any $\tau \in \left(0, \widehat{t}+T\right)$, we have that $\frac{\left(1+\varphi w_{L\tau}^{\epsilon-1}\right)^\mu - 1}{G_t\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^\mu + 1} = \mu\varphi w_{L\tau}^{\epsilon-1} + o\left(\varphi w_{L\tau}^{\epsilon-1}\right)$.[38] Then for any $t \in \left(0, \widehat{t}\right)$:

$$\left(\widehat{h}_t^A\right)^{1-\kappa} \leq \kappa\eta G_t^{\widetilde{\kappa}} \left(\begin{array}{c} \int_t^{\widehat{t}+T} \psi H_t^P \left(\mu\varphi w_{L\tau}^{\epsilon-1} + o\left(\varphi w_{L\tau}^{\epsilon-1}\right)\right) \exp\left(\int_t^\tau \left(g_u^{\pi^N} - r_u\right) du\right) d\tau \\ + \int_{\widehat{t}+T}^\infty \psi H_t^P \frac{\left(1+\varphi w_{L\tau}^{\epsilon-1}\right)^\mu - 1}{G_t\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^\mu + 1} \exp\left(\int_t^\tau \left(g_u^{\pi^N} - r_u\right) du\right) d\tau \end{array}\right).$$

Further, $r_u = \rho + \theta g_u^C$ with $\theta \geq 1$. In addition $C_u = Y_u - X_u$, with X_u the aggregate spending on machines (initially negligible and later on a share of output bounded away from 1) and π_u^N initially grows like Y_u/N_u (and from then on will grow slower), therefore $r_u - g_u^{\pi^N} > \rho$. Hence one can write:

$$\left(\widehat{h}_t^A\right)^{1-\kappa} \leq \kappa\eta G_t^{\widetilde{\kappa}} \left(\int_t^{\widehat{t}+T} \mu\psi H_t^P \varphi w_{L\tau}^{\epsilon-1} e^{\int_t^\tau \left(g_u^{\pi^N} - r_u\right) du} d\tau + o\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right) + o\left(e^{-\rho\left(T+\widehat{t}-t\right)}\right)\right).$$

Since $r_u - g_u^{\pi^N} > \rho$, then $\int_t^{\widehat{t}+T} e^{\int_t^\tau \left(g_u^{\pi^N} - r_u\right) du} d\tau \leq \frac{1}{\rho}\left(1 - e^{-\rho\left(\widehat{t}+T-t\right)}\right)$ and we have:

$$\left(\widehat{h}_t^A\right)^{1-\kappa} \leq \kappa\eta G_t^{\widetilde{\kappa}} \left(\frac{\mu\psi H_t^P \varphi}{\rho} \max_{\tau \in (t, \widehat{t}+T)} \left(w_{L\tau}^{\epsilon-1}\right) + o\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right) + o\left(e^{-\rho T}\right)\right),$$

Therefore, since T is large and $\varphi w_{L\tau}^{\epsilon-1}$ is small, then $\widehat{h}_t^A$ must be small too. In fact, we

[38] The notation $o(z)$ denotes negligible relative to z (that is $f(z) = o(z)$, if $\lim_{z\to 0} f(z)/z = 0$) and $O(z)$ will denote of the same order or negligible in front of z ($f(z) = O(z)$ if $\limsup_{z\to 0} |f(z)/z| < \infty$).

get that $\widehat{h}_t^A = O\left(\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right)^{\frac{1}{1-\kappa}}\right) + o\left(e^{-\rho T}\right)$.

For any $t \in \left(0, \widehat{t}\right)$, we can then rewrite (38) as

$$\frac{\dot{\chi}_t}{\chi_t} = \gamma\psi H^P - \rho - (\theta\psi - \psi + 1)\, g_t^N + O\left(\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right)^{\frac{1}{1-\kappa}}\right) + o\left(e^{-\rho T}\right). \tag{56}$$

Next (4) and the corresponding equation for high-skill labor demand in production imply:

$$\frac{L^{NA}}{L^A} = \frac{(1-G_t)\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-\mu-1}}{G_t} \text{ and } \frac{H^{P,NA}}{H^{P,A}} = \frac{(1-G_t)\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-\mu}}{G_t}.$$

Using these expressions in (5), and knowing that $w_{Lt} = O\left(Y_t/L\right)$ so that $\varphi^{\frac{1}{\epsilon}} Y_t/L = O\left(\varphi^{\frac{1}{\epsilon}} w_{Lt}\right)$, we get:

$$Y_t = \left(1 + O\left(\varphi w_{Lt}^{\epsilon-1}\right)\right) N^{\frac{1}{\sigma-1}} L^{\beta} \left(H_t^P\right)^{1-\beta}. \tag{57}$$

One then gets that wages obey:

$$w_{Ht} = \left(1 + O\left(\varphi w_{Lt}^{\epsilon-1}\right)\right) \frac{\sigma-1}{\sigma} (1-\beta)\, N_t^{\frac{1}{\sigma-1}} L^{\beta} \left(H_t^P\right)^{-\beta}, \tag{58}$$

$$w_{Lt} = \left(1 + O\left(\varphi w_{Lt}^{\epsilon-1}\right)\right) \frac{\sigma-1}{\sigma} \beta N_t^{\frac{1}{\sigma-1}} L^{\beta-1} \left(H_t^P\right)^{1-\beta}. \tag{59}$$

Using (44) we obtain $C_t = Y_t - X_t = \left(1 + O\left(G_t \varphi w_{Lt}^{\epsilon-1}\right)\right) Y_t$, then using the definition of χ_t, (57) and (58), we get:

$$\chi_t = \left(1 + O\left(\varphi w_{Lt}^{\epsilon-1}\right)\right) \sigma\psi L^{\beta(\theta-1)} \left(H_t^P\right)^{(1-\beta)\theta+\beta} N_t^{\frac{(1-\theta)\beta}{(\sigma-1)(1-\beta)}}.$$

Differentiating and plugging into (56) and using (46), we get (recalling (59) so that $d\ln\left(1 + O\left(\varphi w_{Lt}^{\epsilon-1}\right)\right)/dt$ will be of order $O\left(\varphi w_{Lt}^{\epsilon-1}\right)$ as well).

$$\left((1-\beta)\,\theta + \beta\right) \frac{\dot{H}_t^P}{H_t^P} = \gamma\psi H_t^P - \rho - \left(\frac{\theta-1}{\sigma-1} + 1\right) \gamma \left(H - H_t^P\right) + O\left(\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right)^{\frac{1}{1-\kappa}}\right) + o\left(e^{-\rho T}\right), \tag{60}$$

we dropped terms in $\varphi w_{Lt}^{\epsilon-1}$ since there will negligible in front of $\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right)^{\frac{1}{1-\kappa}}$. The exact counterpart of this system admits a BGP with H_t^P constant without transitional

dynamics as in Romer (1990). Therefore, we have over the interval $\left(0,\widehat{t}\right)$

$$H_t^P = \frac{\left(\frac{\theta-1}{\sigma-1}+1\right)H+\frac{\rho}{\gamma}}{\psi+\frac{\theta-1}{\sigma-1}+1} + O\left(\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right)^{\frac{1}{1-\kappa}}\right) + o\left(e^{-\rho T}\right)$$
$$\text{and } g_t^N = \frac{\gamma H\psi-\rho}{\psi+\frac{\theta-1}{\sigma-1}+1} + O\left(\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right)^{\frac{1}{1-\kappa}}\right) + o\left(e^{-\rho T}\right) \tag{61}$$

which is positive under assumption (48). With a low $\widehat{h}_t^A$, (12) can be solved as $G_t = G_0 \exp\left(-\frac{\gamma H\psi-\rho}{\psi+\frac{\theta-1}{\sigma-1}+1}t\right) + O\left(\left(\varphi w_{L\widehat{t}+T}^{\epsilon-1}\right)^{\frac{1}{1-\kappa}}\right) + o\left(e^{-\rho T}\right)$. This characterizes the solution over the interval $\left(0,\widehat{t}\right)$ for w_{Lt} sufficiently small relative to $\widetilde{\varphi}^{-1}$. A small w_{Lt} in return can be generated by a sufficiently small N_0, so that for any $a > 0$, there is a N_0 low enough that the automation rate is lower than a. This establishes Proposition 3.B.

6.3.4 Proof of Proposition 3.C

We now prove Part C. The proof relies on two lemmas proved in Appendix 7.5.3.

Lemma 2. *If $\kappa(1-\beta)+\widetilde{\kappa} < 1$, and g_t^N has a positive limit then G_t cannot converge toward 0.*

This lemma shows that automation must take off at some point, the second lemma is more technical:

Lemma 3. *If G_t is bounded above 0 then $\hat{h}_t^A$ is bounded.*

Under the assumptions of the Proposition, G_t has a limit G_∞ and g_t^N a positive limit g_∞^N. Then Lemma 2 implies that $G_\infty > 0$ and Lemma 3 together with (31) that $G_\infty < 1$. Following Proposition (2), we then get that $w_{Lt} = O\left(N_t^{\frac{\psi}{1+\beta(\sigma-1)}}\right)$ or $\omega_t = O(1)$. Therefore, we can rewrite the system as (31),

$$\dot{\hat{h}}_t^A = \frac{\gamma\left(\widehat{h}_t^A\right)^2}{\kappa} + \eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^\kappa\left(\widehat{h}_t^A - \frac{\kappa\hat{\pi}_t^A}{(1-\kappa)\hat{v}_t}\right) + \frac{\widetilde{\kappa}\left(\eta G_t^{\widetilde{\kappa}-1}\left(\widehat{h}_t^A\right)^{\kappa+1}(1-G_t) - g_t^N\hat{h}_t^A\right)}{1-\kappa} + O(n_t),$$

$$\dot{\chi}_t = \chi_t\left(\gamma\frac{1-\kappa}{\kappa}\widehat{h}_t^A - \rho - (\theta\psi-\psi+1)g_t^N\right) + O(n_t).$$

Parameter	σ	ϵ	β_1	γ	$\tilde{\kappa}$	θ	η	κ	ρ	β_2	Δ	β_4	$\tilde{\varphi}$	N_{1963}	G_{1963}
Value	5.80	4.34	0.60	0.43	0.55	1	0.40	0.66	0.034	0.16	0.011	0.76	1.56	14.8	0.03

Table 2: Parameters (only matching the first 30 years)

Knowing that

$$H_t^P = \left(\frac{\sigma-1}{\sigma}\right)^{\frac{1}{1-\beta}\left(\frac{1}{\theta}-\beta\right)} \frac{(1-\beta)^{\frac{1}{\theta}}\,\beta^{\frac{\beta}{1-\beta}\left(\frac{1}{\theta}-1\right)}\chi_t^{\frac{1}{\theta}}\left(G_t\varphi^{\mu}\right)^{\psi\left(\frac{1}{\theta}-1\right)}}{G_t\varphi^{\mu}\left(1-\beta\frac{\sigma-1}{\sigma}\right)} + O\left(n_t^{\frac{1}{\mu}}\right), \tag{62}$$

$$\frac{\widehat{\pi}_t^A}{\widehat{v}_t} = \frac{\psi H_t^P}{G_t} + O\left(n_t^{\frac{1}{\mu}}\right), \tag{63}$$

and (46). Using that g_t^N and G_t have limits in (31) implies that $\hat{h}_t^A$ must also have a limit. Using (46), this implies that H_t^P must also have a limit and therefore using (62) that χ_t must have a limit. In other words, the equilibrium path tends toward the steady-state $\left(\widehat{h}^{A*}, G^*, \chi^*\right)$ with $n_t \to 0$ defined in Lemma 1.

6.4 Appendix to the Quantitative Exercise

This Appendix presents two exercises which complement Section 4: a calibration on the first 30 years of data and an analysis of the role played by the automation externality. Appendix 7.11 contains additional details on the quantitative exercise.

6.4.1 Out-of-sample prediction

We reproduce our calibration exercise but only matching the first 30 years of data. Figure 8 reports the results and Table 2 gives the new parameters. The model behaves very well out of sample. The predicted path and parameters are close to those of the baseline case. The calibrated model over the first 30 years slightly underestimates the pace of the rise of the skill-premium. The biggest parameter difference is a lower elasticity of substitution between low-skill workers and machines at 4.3 instead of 6. This decreases the incentive to automate, slows the growth in the skill-premium and lowers the growth rate of equipment (Panel D).

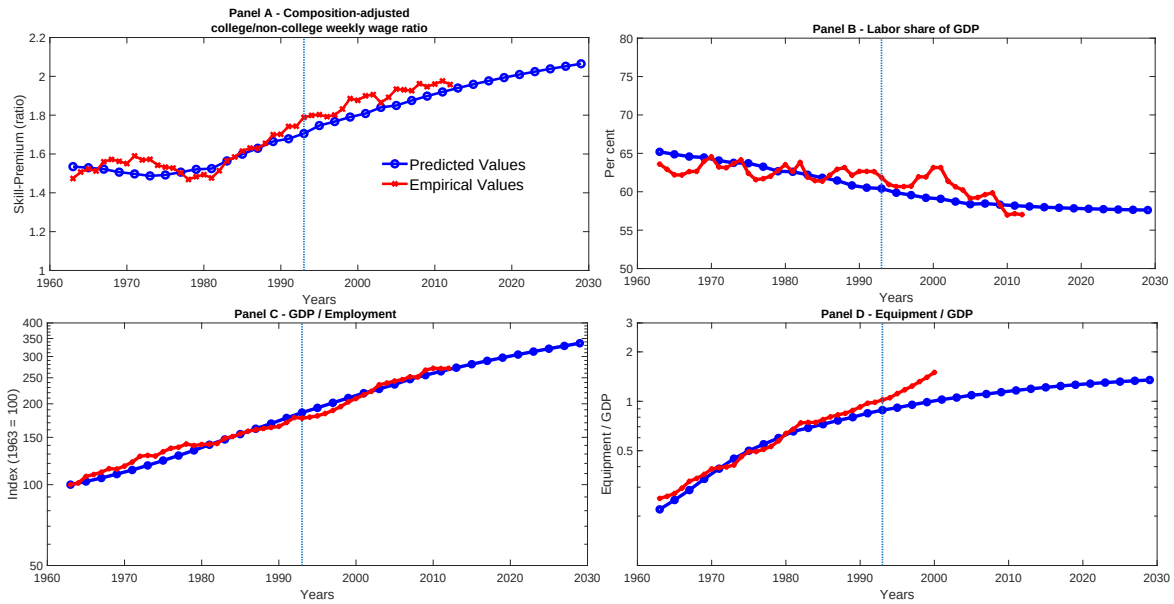

Figure 8: Predicted and empirical time paths only matching the first 30 years (until 1992)

Parameter	σ	ϵ	β_1	γ	$\tilde{\kappa}$	θ	η	κ	ρ	β_2	Δ	β_4	$\tilde{\varphi}$	N_{1963}	G_{1963}
Value	4.95	4.95	0.66	0.51	0	1.15	0.30	0.87	0.039	0.14	0.011	0.76	1.47	18.4	0.12

Table 3: Parameters (when $\tilde{\kappa}$=0)

6.4.2 The role of the automation externality

To analyze the role played by the automation externality, we recalibrate our model without it (i.e., we impose that $\tilde{\kappa} = 0$). Table 3 gives the resulting parameters and Figure 9 reports the results. The model still reproduces the paths for the labor share, GDP/employment and equipment/GDP. Yet, it does not capture the evolution of the skill premium. Indeed, the fast rise in the skill premium in the 1980s and 1990s requires a fast increase in automation which, given the moderate decline in the labor share and the stable economic growth, can only be brought about by a positive automation externality. The data clearly favor a positive automation externality (even though the exact value of $\tilde{\kappa}$ is not precisely estimated, see Table 7).

Furthermore, we reproduce the exercise of Section 4.2 without an automation externality. We fix the parameters to their values in the recalibrated model with $\tilde{\kappa} = 0$, and look for the (exogenous) paths of N_t and G_t that minimize the distance between the model and the data. Figure 10 shows that with this set of parameters, an exogenous technological model would indeed require a faster increase in G_t to better match the data than in the endogenous growth model: the path looks closer to that obtained in

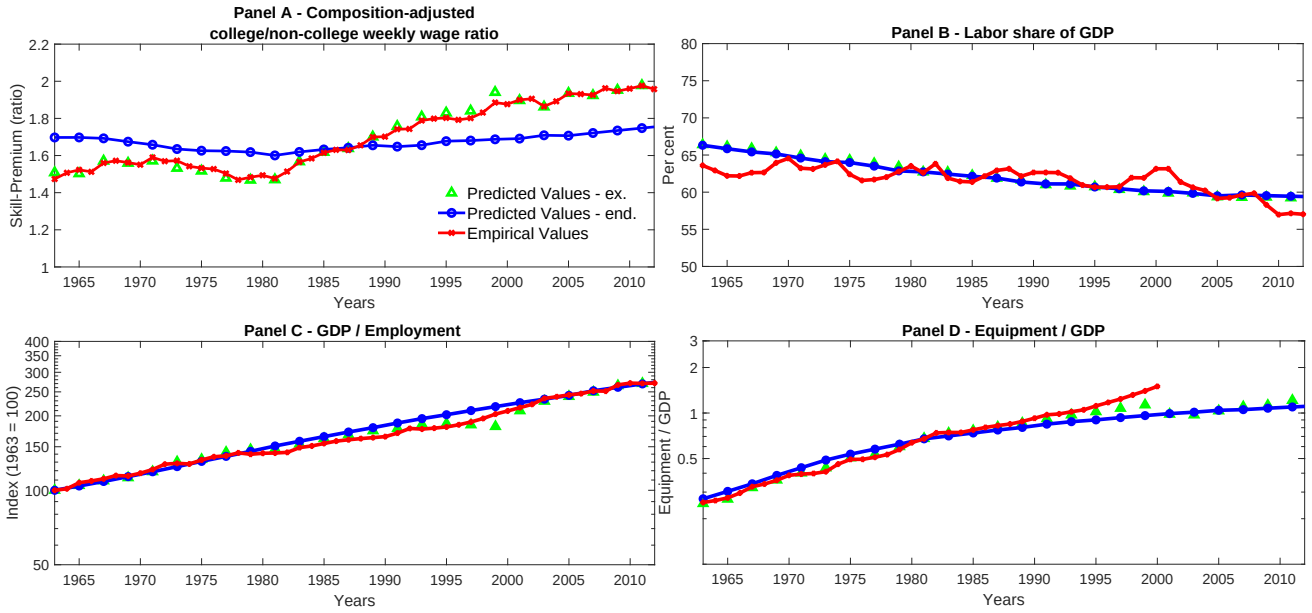

Figure 9: Empirical paths and predicted paths for the endogenous and exogenous technology models for $\tilde{\kappa} = 0$

the endogenous model with $\tilde{\kappa} \neq 0$ than with $\tilde{\kappa} = 0$. If anything, it exhibits an even sharper increase in the 1980s and 1990s than in Figure5.B. Figure 9 shows that this exogenous growth model still matches the data well.

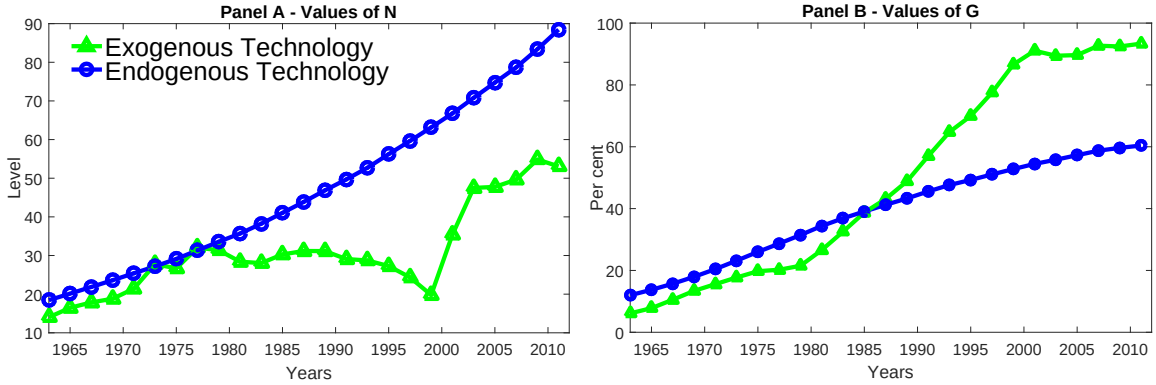

Figure 10: Paths for N_t and G_t in the endogenous and exogenous growth models without automation externality.

7 Secondary Appendix (For Online Publication)

7.1 Additional results for Section 2

7.1.1 Proof of Proposition 1 in the perfect substitute case: $\epsilon = \infty$

In the perfect substitute case, there are three possibilities. Case i) $w_L < \widetilde{\varphi}^{-1}$: automated firms only use low-skill workers and low-skill wages are given by

$$w_L = \frac{\sigma - 1}{\sigma} \beta \left(\frac{H^P}{L} \right)^{1-\beta} N^{\frac{1}{\sigma-1}}, \tag{64}$$

with a skill premium obeying $\frac{w_H}{w_L} = \frac{1-\beta}{\beta} \frac{L}{H^P}$.

Case ii) $w_L = \widetilde{\varphi}^{-1}$: automated firms use machines but also possibly workers, in which case high-skill wages can be obtained from (6) which is now written as:

$$\frac{\sigma}{\sigma - 1} \frac{N^{\frac{1}{1-\sigma}}}{\beta^{\beta} (1-\beta)^{1-\beta}} \widetilde{\varphi}^{-\beta} w_H^{1-\beta} = 1.$$

Case iii) $w_L > \widetilde{\varphi}^{-1}$ and all automated firms use machines only, in that case, we get that (25) is replaced by

$$w_L = \frac{\sigma - 1}{\sigma} \beta \left(\frac{H^P}{L} \right)^{(1-\beta)} N^{\frac{1}{\sigma-1}} (1-G)^{1-\beta} \left(G \left(\widetilde{\varphi} w_L \right)^{\beta(\sigma-1)} + 1 - G \right)^{\frac{1}{\sigma-1} - (1-\beta)}, \tag{65}$$

and the skill premium obeys:

$$\frac{w_H}{w_L} = \frac{1-\beta}{\beta} \frac{L}{H^P} \frac{G \left(w_L \widetilde{\varphi} \right)^{\beta(\sigma-1)} + 1 - G}{1 - G}. \tag{66}$$

One can rewrite (65) as

$$w_L^{1-\beta} = \frac{\sigma - 1}{\sigma} \beta \left(\frac{H^P}{L} \right)^{(1-\beta)} \frac{N^{\frac{1}{\sigma-1}} (1-G)^{1-\beta} \left(G + (1-G) \left(\widetilde{\varphi} w_L \right)^{-\beta(\sigma-1)} \right)^{\frac{1}{\sigma-1}} \widetilde{\varphi}^{\beta}}{\left(G \left(\widetilde{\varphi} w_L \right)^{\beta(\sigma-1)} + 1 - G \right)^{1-\beta}}.$$

The left-hand side increases in w_L and the right-hand side decreases in w_L, so this

expression defines w_L uniquely. Moreover w_L is greater than $\widetilde{\varphi}^{-1}$ if and only if

$$N^{\frac{1}{\sigma-1}}(1-G)^{1-\beta} > \frac{\sigma}{(\sigma-1)\beta\widetilde{\varphi}}\left(\frac{L}{H^P}\right)^{(1-\beta)}.$$

Hence w_L and w_H are defined uniquely as functions of N, G and H^P. If $N^{\frac{1}{\sigma-1}} < \frac{\sigma}{(\sigma-1)\beta\widetilde{\varphi}}\left(\frac{L}{H^P}\right)^{(1-\beta)}$, we are in case i), if $N^{\frac{1}{\sigma-1}}(1-G)^{1-\beta} \leq \frac{\sigma}{(\sigma-1)\beta\widetilde{\varphi}}\left(\frac{L}{H^P}\right)^{(1-\beta)} \leq N^{\frac{1}{\sigma-1}}$ then we are in case ii) and if $N^{\frac{1}{\sigma-1}}(1-G)^{1-\beta} > \frac{\sigma}{(\sigma-1)\beta\widetilde{\varphi}}\left(\frac{L}{H^P}\right)^{(1-\beta)}$, we are in case iii).

It is then direct to show that w_H increases in N and weakly increases in G, that w_H/w_L is weakly increasing in N and G (weakly because of case i)), and that w_L is weakly increasing in N (weakly because of case ii)).

Comparative statics of w_L with respect to G. Furthermore, (65) shows that w_L is decreasing in G in case iii) if $\frac{1}{\sigma-1} \leq 1-\beta$. Therefore w_L is weakly decreasing in G if $\frac{1}{\sigma-1} \leq 1-\beta$.

Assume now that $\frac{1}{\sigma-1} > 1-\beta$. Log-differentiating (65), one gets:

$$\widehat{w}_L = \left[\left(\frac{1}{\sigma-1} - (1-\beta)\right)\frac{\left((\widetilde{\varphi}w_L)^{\beta(\sigma-1)} - 1\right)G}{G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)} + (1-G)} - (1-\beta)\frac{G}{1-G}\right]\frac{\widehat{G}}{Den} \tag{67}$$

where

$$Den \equiv 1 - \beta\frac{G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)}}{G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)} + 1 - G} + \frac{(1-\beta)\beta(\sigma-1)G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)}}{G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)} + 1 - G}. \tag{68}$$

We have

$$\begin{aligned} & \left(\frac{1}{\sigma-1} - (1-\beta)\right)\frac{\left((\widetilde{\varphi}w_L)^{\beta(\sigma-1)} - 1\right)G}{G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)} + 1 - G} - (1-\beta)\frac{G}{1-G} \\ < \; & \frac{1}{\sigma-1} - \frac{1-\beta}{1-G}, \end{aligned}$$

which is negative for G large enough. Hence we obtain that for G high enough, w_L is weakly decreasing in G (strictly in case iii)).

Increase in the number of non-automated products. In case i) an increase in the mass of non-automated products leads to an increase in w_H and w_L while w_H/w_L is constant. In case ii), w_H increases, w_H/w_L increases and w_L is constant.

Log-differentiating (65) with respect to both N and G, one gets:

$$\widehat{w}_L = \left[\left(\left(\frac{1}{\sigma-1} - (1-\beta)\right)\frac{\left((\widetilde{\varphi}w_L)^{\beta(\sigma-1)} - 1\right)G}{G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)} + 1 - G} - (1-\beta)\frac{G}{1-G}\right)\widehat{G} + \frac{1}{\sigma-1}\widehat{N}\right]\frac{1}{Den}.$$

For an increase in the mass of non-automated products, $\widehat{G} = -\widehat{N}$, so that the change in w_L in that case is given by:

$$\widehat{w}_L^{NT} = \left[(1-\beta)\left(\frac{\left((\widetilde{\varphi}w_L)^{\beta(\sigma-1)} - 1\right)G}{G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)} + 1 - G} + \frac{G}{1-G}\right) + \frac{1}{\sigma-1}\left(1 - \frac{\left((\widetilde{\varphi}w_L)^{\beta(\sigma-1)} - 1\right)G}{G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)} + 1 - G}\right)\right]\frac{\widehat{N}}{Den}.$$

Therefore w_L increases.

Log-differentiating (66), we get:

$$\widehat{w}_H = \frac{G\left((\widetilde{\varphi}w_L)^{\beta(\sigma-1)} - 1\right)}{G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)} + 1 - G}\widehat{G} + \frac{G}{1-G}\widehat{G} + \left(1 + \frac{\beta(\sigma-1)G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)}}{G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)} + 1 - G}\right)\widehat{w}_L.$$

Therefore, for an increase in the mass of non-automated products, one gets:

$$\widehat{w}_H^{NT} = \frac{1}{\sigma-1}\left(1 - \frac{\left((\widetilde{\varphi}w_L)^{\beta(\sigma-1)} - 1\right)G}{G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)} + 1 - G}\right)\frac{\widehat{N}}{Den},$$

which ensures that w_H increases with the mass of non-automated products. Finally,

$$\widehat{w}_H^{NT} - \widehat{w}_L^{NT} = -\frac{(1-\beta)G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)}}{\left(G\left(\widetilde{\varphi}w_L\right)^{\beta(\sigma-1)} + 1 - G\right)(1-G)}\frac{\widehat{N}}{Den}.$$

Therefore an increase in the mass of non-automated products decreases the skill premium in case iii).

Overall we get that an increase in the mass of non-automated products weakly increases w_L, increases w_H and decreases w_H/w_L if N is large enough but $G \neq 1$ (so that we are in case iii)).

7.1.2 Comparative statics with respect to $\widetilde{\varphi}$.

We now look at the effect of an increase in machine's productivity $\widetilde{\varphi}$ (which up to some relabeling is equivalent to a decline in the price of machines). We focus on the case $\epsilon < \infty$, so that an increase in $\widetilde{\varphi}$ is equivalent to an increase in φ. To look at its effect on low-skill wages, log-differentiate (25)

$$\widehat{w}_L = \frac{Gw^{\varepsilon-1}\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu-1}}{Den\left(1+\varphi w_L^{\varepsilon-1}\right)}\left(\frac{(1-\beta)(\mu-1)}{G\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu-1}+1-G}+\frac{\mu\left(\frac{1}{\sigma-1}-(1-\beta)\right)\left(1+\varphi w_L^{\varepsilon-1}\right)}{G\left(1+\varphi w_l^{\varepsilon-1}\right)^{\mu}+1-G}\right)\widehat{\varphi},$$

where Den is still given by (68). We then get that $\frac{\partial w_L}{\partial \varphi} < 0$ if $\psi \leq 1$ (that is $(1-\beta)(\sigma-1) > 1$), in which case $\frac{\partial w_L}{\partial G} < 0$. Provided that $\psi > 1$, we have that

$$\frac{\partial w_L}{\partial \varphi} < 0 \Leftrightarrow \frac{G\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu}+1-G}{G\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu-1}+1-G}\frac{1}{\psi-1} > \frac{\mu\left(1+\varphi w_L^{\varepsilon-1}\right)}{1-\mu}.$$

Using (26), we get

$$\frac{\partial w_L}{\partial G} < 0 \Leftrightarrow \frac{G\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu}+1-G}{G\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu-1}+1-G}\frac{1}{\psi-1} > \frac{\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu}-1}{1-\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu-1}}.$$

Note that

$$\frac{\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu}-1}{1-\left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu-1}} < \frac{\mu\left(1+\varphi w_L^{\varepsilon-1}\right)}{1-\mu} \Leftrightarrow \left(1+\varphi w_L^{\varepsilon-1}\right)^{\mu} < 1+\mu\varphi w_L^{\varepsilon-1},$$

which is always true since $\mu < 1$. Therefore, $\frac{\partial w_L}{\partial \varphi} < 0$ implies that $\frac{\partial w_L}{\partial G} < 0$ but the reverse is not true.

7.2 Proofs of the asymptotic results

7.2.1 Asymptotic results when $G_\infty \notin (0,1)$

In this subsection, we extend Proposition 2 to the cases where $G_\infty \notin (0,1)$.

Proposition 4. *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 limits G_∞, g_∞^N and H_∞^P with $g_\infty^N > 0$ and $H_\infty^P > 0$.*

A). If $G_\infty = 1$, the asymptotic growth rates of w_{Ht} and Y_t also obey (10). If G_t converges sufficiently fast (such that $\lim_{t\to\infty} (1-G_t)\, N_t^{\psi(1-\mu)\frac{\epsilon-1}{\epsilon}}$ exists and is finite) then :

-i) If $\epsilon < \infty$ the asymptotic growth of w_{Lt} is positive at :

$$g_\infty^{w_L} = g_\infty^Y/\epsilon. \tag{69}$$

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

B) If $G_\infty = 0$ and G_t converges sufficiently fast (such that $\lim_{t\to\infty} G_t N_t^\beta$ exists and is finite), then the asymptotic growth rates of w_{Lt}, w_{Ht} and Y_t obey:

$$g_\infty^{w_L} = g_\infty^{w_H} = g_\infty^Y = g_\infty^N/(\sigma-1). \tag{70}$$

Proof. **Case where $G_\infty = 1$ (Part A)**. The proof of Proposition 2 directly applies to show that the asymptotic growth rates of w_{Ht} and Y_t also obey (10).

Subcase with $\epsilon < \infty$. With $G_\infty = 1$, equation (25) still implies that w_{Lt} is unbounded and gives:

$$w_{Lt} \sim \left(\left(\frac{\sigma-1}{\sigma}\beta\right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L}\varphi^{\mu(\psi-1)}\right)^{\frac{1}{\epsilon}} N_t^{\frac{\psi}{\epsilon}} \left(\varphi^{\mu-1} + (1-G_t)\, w_{Lt}^{(\epsilon-1)(1-\mu)}\right)^{\frac{1}{\epsilon}}. \tag{71}$$

Following the assumption of Part A in Proposition 4, we assume that $\lim (1-G_t)\, N_t^{\frac{\psi}{\epsilon}(\epsilon-1)(1-\mu)}$ exists and is finite. Suppose first that $\limsup (1-G_t)\, w_{Lt}^{(\epsilon-1)(1-\mu)} = \infty$, then there must exist a sequence of t's, denoted t_n for which:

$$w_{Lt_n} \sim \left(\left(\frac{\sigma-1}{\sigma}\beta\right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L}\left(\varphi^\mu\right)^{\psi-1}\right)^{\frac{1}{1+\beta(\sigma-1)}} \left((1-G_{t_n})\, N_{t_n}^\psi\right)^{\frac{1}{1+\beta(\sigma-1)}}.$$

Yet, this implies

$$(1-G_{t_n})\, w_{Lt_n}^{(\epsilon-1)(1-\mu)} \sim \begin{array}{l} \left(\left(\frac{\sigma-1}{\sigma}\beta\right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L}\left(\varphi^\mu\right)^{\psi-1}\right)^{\frac{(\epsilon-1)(1-\mu)}{1+\beta(\sigma-1)}} \\ \times \left((1-G_{t_n})\, N_{t_n}^{\frac{\psi(\epsilon-1)(1-\mu)}{\epsilon}}\right)^{\frac{\epsilon}{1+\beta(\sigma-1)}} \end{array},$$

the left-hand side is assumed to be unbounded, while the right-hand side is bounded: there is a contradiction. Therefore, $\limsup (1-G_t)\, w_{Lt}^{(\epsilon-1)(1-\mu)} < \infty$.

Consider now the possibility that $\lim (1-G_t)\, w_{Lt}^{(\epsilon-1)(1-\mu)} = 0$, then (71) implies

$$w_{Lt} \sim \left(\left(\frac{\sigma-1}{\sigma}\beta \right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L} \varphi^{\mu\psi-1} \right)^{\frac{1}{\epsilon}} N_t^{\frac{\psi}{\epsilon}}.$$

Therefore we get that $g_\infty^{w_L} = \frac{\psi}{\epsilon} g_\infty^N = \frac{1}{\epsilon} g_\infty^Y$.[39]

Alternatively, $\limsup (1-G_t)\, w_{Lt}^{(\epsilon-1)(1-\mu)}$ is finite but strictly positive (given by λ_1). In this case, there exists a sequence of $t's$, denoted t_m such that

$$w_{Lt_m} \sim \left(\left(\frac{\sigma-1}{\sigma}\beta \right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L} \varphi^{\mu(\psi-1)} \left(\varphi^{\mu-1} + \lambda_1 \right) \right)^{\frac{1}{\epsilon}} N_{t_m}^{\frac{\psi}{\epsilon}}. \tag{72}$$

This leads to

$$\lambda_1 \sim \left(\left(\frac{\sigma-1}{\sigma}\beta \right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L} \varphi^{\mu(\psi-1)} \left(\varphi^{\mu-1} + \lambda_1 \right) \right)^{\frac{(\epsilon-1)(1-\mu)}{\epsilon}} (1-G_{t_m})\, N_{t_m}^{\frac{\psi}{\epsilon}(\epsilon-1)(1-\mu)},$$

which is only possible if $\lim (1-G_t)\, N_t^{\frac{\psi}{\epsilon}(\epsilon-1)(1-\mu)} > 0$. We denote such a limit by λ. Then (71) leads to

$$\left(w_{Lt}^\epsilon N_t^{-\psi} \right) \sim \left(\frac{\sigma-1}{\sigma}\beta \right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L} \varphi^{\mu(\psi-1)} \left(\varphi^{\mu-1} + \lambda \left(N_t^{-\psi} w_{Lt}^\epsilon \right)^{\frac{(\epsilon-1)(1-\mu)}{\epsilon}} \right),$$

which defines uniquely the limit of $w_{Lt}^\epsilon N_t^{-\psi}$. We then obtain that $g_\infty^{w_L} = \frac{\psi}{\epsilon} g_\infty^N$. This completes the poof of part A.

Subcase with $\epsilon = \infty$. Low skill wages are defined as described in Appendix 7.1.1. With $G_\infty = 1$ and knowing that $\lim (1-G_t)\, N_t^\psi$ exists and is finite, (64) implies that w_{Lt} must be bounded weakly above $\widetilde{\varphi}$ in the long-run. As a result, (65) leads to

$$w_{Lt} \sim \left(\left(\frac{\sigma-1}{\sigma}\beta \widetilde{\varphi}^{\beta\left(1-\psi^{-1}\right)} \right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L} \right)^{\frac{1}{1+\beta(\sigma-1)}} \left((1-G_t)\, N_t^\psi \right)^{\frac{1}{1+\beta(\sigma-1)}} \text{ if } w_{Lt} > \widetilde{\varphi}.$$

Since $\lim (1-G_t)\, N_t^\psi$ exists and is finite, w_{Lt} also admits a finite limit. In particular, if $\lim (1-G_t)\, N_t^\psi = 0$, then $w_{L\infty} = \widetilde{\varphi}$.

[39] Expressions regarding the asymptotic growth rates (here and below) assume existence of the limits but expressions on equivalence ($\sim$) or orders of magnitude (O) do not.

Case where $G_\infty = 0$ (Part B). If $\lim G_t = 0$ then (25) implies that for $\epsilon < \infty$:

$$w_{Lt} \sim \frac{\sigma-1}{\sigma}\beta\left(\frac{H_\infty^P}{L}\right)^{(1-\beta)} N_t^{\frac{1}{\sigma-1}} \left(G_t\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^\mu + 1\right)^{\frac{1}{\sigma-1}-(1-\beta)}.$$

This expression directly implies that $\lim w_{Lt} = \infty$ (otherwise there is a subsequence where the left-hand side is bounded while the right-hand side is unbounded). Therefore we actually get:

$$w_{Lt} \sim \frac{\sigma-1}{\sigma}\beta\left(\frac{H_\infty^P}{L}\right)^{(1-\beta)} N_t^{\frac{1}{\sigma-1}} \left(G_t w_{Lt}^{\beta(\sigma-1)}\varphi^\mu + 1\right)^{\frac{1}{\sigma-1}-(1-\beta)}. \tag{73}$$

Note that if $\epsilon = \infty$, then we must be in case iii) when $G_\infty = 0$ and (65) also directly implies (73) (as $\varphi^\mu = \widetilde{\varphi}^{\beta(\sigma-1)}$ in that case).

Assume that $\lim_{t\to\infty} G_t N_t^\beta = \lambda$ exists and is finite. Then (73) implies:

$$w_{Lt}N_t^{-\frac{1}{\sigma-1}} \sim \frac{\sigma-1}{\sigma}\beta\left(\frac{H_\infty^P}{L}\right)^{(1-\beta)} \left(\lambda\left(w_{Lt}N_t^{-\frac{1}{\sigma-1}}\right)^{\beta(\sigma-1)}\varphi^\mu + 1\right)^{\frac{1}{\sigma-1}-(1-\beta)},$$

which implies that $\lim_{t\to\infty} w_{Lt}N_t^{-\frac{1}{\sigma-1}}$ exists and is finite as well. Therefore one gets that $g_\infty^{w_L} = g_\infty^N/(\sigma-1)$. Using (24) then immediately implies (70). □

Part A of the proposition shows that when $G_\infty = 1$, w_{Lt} is bounded when there is economy-wide perfect substitution (that is we also have $\epsilon = \infty$), even then, low-skill wages are bounded below by $\tilde{\varphi}^{-1}$, as a lower wage would imply that no firm would use machines. If instead $\epsilon < \infty$, then low-skill wages must grow asymptotically (similarly to the case $G_\infty < 1$), but low-skill workers now derive their income asymptotically from automated firms and the asymptotic growth rate depends on the elasticity of substitution between machines and low-skill workers in automated firms, ϵ.

Part B of the proposition shows that when the share of automated products converges toward 0 sufficiently fast, the economy behaves like in a classic expanding-variety model and low-skill and high-skill wages grow at the same rate.

7.2.2 Sufficient conditions for $G_\infty \in (0,1)$.

We prove the following Lemma:

Lemma 4. *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* $0 < G_\infty < 1$.

Note that $G_t N_t$ is the mass of automated firms and let $\nu_{1,t} > 0$ be the intensity at which non-automated firms are automated at time t and $0 \leq \nu_{2,t} < 1$ be the fraction of new products introduced at time t that are initially automated. Then $\dot{(G_t N_t)} = \nu_{1,t}(1 - G_t)N_t + \nu_{2,t}\dot{N}_t$ such that $\dot{G}_t = \nu_{1,t}(1 - G_t) - (G_t - \nu_{2,t})g_t^N$. First assume that $G_\infty = 1$, then if $\nu_{1,t} < \bar{\nu}_1 < \infty$ and $\nu_{2,t} < \bar{\nu}_2 < 1$, we get that $\dot{G}_t$ must be negative for sufficiently large t, which contradicts the assumption that $G_\infty = 1$. Similarly if $G_\infty = 0$, then having $\nu_{1,t} > \underline{\nu}$ for all t, gives that $\dot{G}_t$ must be positive for sufficiently large t, which also implies a contradiction. Hence a limit must have $0 < G_\infty < 1$.

7.3 An endogenous supply response in the skill distribution: static model

We present here an extension of the baseline model with 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 $q \to \infty$ implying equal distribution of skills and $q < \infty$ implying a ranking of increasing endowments of high-skill on $[0, \bar{H}(1 + q)/q]$.

The supply of low-skill and high-skill labor are now endogenous. This does not affect (6) which still holds. (7) also holds with L_t replacing L and knowing that H_t^P obeys (13) but with H_t instead of H in the right-hand side. Because workers are ordered such that a worker with a higher index j supplies relatively more high-skill labor, then at all point in times there exists a threshold $\bar{j}_t$ such that workers $j \in \left(0, \bar{j}_t\right)$ supply low-skill labor and workers $j \in \left(\bar{j}_t, 1\right)$ supply high-skill labor. As a result, we get that the total mass of low-skill labor is:

$$L_t = l\overline{H}\bar{j}_t, \tag{74}$$

and the mass of high-skill labor is

$$H_t = \overline{H}\left(1 - \overline{j}_t^{\frac{1+q}{q}}\right) \leq \overline{H}. \tag{75}$$

The cut-off $\overline{j}_t$ obeys $l\overline{H}w_{Lt} = \Gamma\left(\overline{j}_t\right) w_{Ht}$, that is

$$\overline{j}_t = \left(\frac{q}{1+q}\frac{lw_{Lt}}{w_{Ht}}\right)^q. \tag{76}$$

$\overline{j}_t$ decreases as the skill premium increases and q measures the elasticity of $\overline{j}_t$ with respect to the skill premium.

A proposition similar to Proposition 2 applies but the asymptotic growth rate of low-skill wages is higher:

Proposition 5. *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 limits G_∞, g_∞^N and H_∞^P with $G_\infty \in (0,1)$, $g_\infty^N > 0$ and $H_\infty^P > 0$. Then the asymptotic growth of high-skill wages w_{Ht}, output Y_t and low-skill wages are:*

$$g_\infty^{w_H} = g_\infty^Y = g_\infty^N \text{ and } g_\infty^{w_L} = \frac{1+q}{1+q+\beta(\sigma-1)} g_\infty^Y. \tag{77}$$

Proof. We consider processes $\left(N_t, G_t, H_t^P\right)$ such that g_t^N, G_t and H_t^P admit strictly positive limits. Plugging (76) and (74) in (7), we get:

$$\frac{w_{Ht}}{w_{Lt}} = l\left(\frac{1-\beta}{\beta}\frac{\overline{H}}{H_t^P}\left(\frac{q}{1+q}\right)^q \frac{G_t + (1-G_t)\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-\mu}}{G_t\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-1} + (1-G_t)\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-\mu}}\right)^{\frac{1}{1+q}}, \tag{78}$$

which together with (6) determines w_{Ht} and w_{Lt} for given $\left(N_t, G_t, H_t^P\right)$. From then on the reasoning follows that of Appendix 6.2. First, we derive that $w_{L\infty} > 0$, such that $g_\infty^{w_H} = g_\infty^{GDP} = \psi g_\infty^N$, and that we must have $g_\infty^{w_L} < g_\infty^{w_H}$, such that $\overline{j}_\infty = 0$. Second, we study the asymptotic behavior of w_{Lt} both when $\epsilon < \infty$ and when $\epsilon = \infty$.

Case with $\epsilon < \infty$. Plugging (78) in (6) gives w_{Lt} in function of N_t, G_t and H_t^P:;

$$w_{Lt} = \begin{array}{c} \frac{\sigma-1}{\sigma}\beta^{\frac{1+\beta q}{1+q}}\left((1-\beta)\frac{1+q}{q}\right)^{\frac{(1-\beta)q}{1+q}} \frac{1}{l^{1-\beta}}\left(\frac{H_t^P}{\overline{H}}\right)^{\frac{1-\beta}{1+q}} N^{\frac{1}{\sigma-1}} \\ \times \frac{\left(G_t\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{\mu-1} + (1-G_t)\right)^{\frac{1-\beta}{1+q}}}{\left(G_t\left(\varphi w_{Lt}^{\epsilon-1}+1\right)^{\mu} + 1 - G_t\right)^{\frac{1}{1-\sigma}+\frac{1-\beta}{1+q}}} \end{array}, \tag{79}$$

which replaces (25). It is direct that when $G_\infty < 1$, we obtain (77). In this case, we further have

$$g_\infty^{\bar{j}} = q\left(g_\infty^{w_L} - g_\infty^{w_H}\right) = -\frac{q\beta\left(\sigma-1\right)}{1+q+\beta\left(\sigma-1\right)} g_\infty^{GDP}. \tag{80}$$

Case with $\epsilon = \infty$. In this case, (79) becomes

$$w_{Lt} = \begin{array}{c} \frac{\sigma-1}{\sigma}\beta\left(\frac{1+q}{q}\right)^{(1-\beta)q}\left(\frac{H_t^P}{l\overline{H}}\right)^{1-\beta} N^{\frac{1}{\sigma-1}}\left(1-G_t\right)^{\frac{1-\beta}{1+q}} \\ \times\left(G_t\left(\widetilde{\varphi}w_{Lt}\right)^{\beta(\sigma-1)} + \left(1-G_t\right)\right)^{\frac{1}{\sigma-1}-\frac{1-\beta}{1+q}} \end{array}, \text{ if } w_{Lt} > \widetilde{\varphi}^{-1},$$

$$w_{Lt} = \frac{\sigma-1}{\sigma}\beta\left(\frac{1+q}{q}\right)^{(1-\beta)q}\left(\frac{H_t^P}{l\overline{H}}\right)^{1-\beta} N^{\frac{1}{\sigma-1}}, \text{ if } w_{Lt} < \widetilde{\varphi}^{-1}.$$

Once again, following the steps of Appendix 6.2, we get that if $G_\infty < 1$, (77) applies (and accordingly we also get (80)). □

Intuitively, as low- and high-skill wages diverge, workers switch from being low-skill to high-skill. This endogenous supply response 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.

7.4 Alternative production technology for machines

The assumption of identical production technologies for consumption and machines imposes a constant real price of machines once they are introduced. As shown in Nordhaus (2007) the price of computing power has dropped dramatically over the past 50 years and the declining real price of computers/capital is central to the theories of Autor and Dorn (2013) and Karabarbounis and Neiman (2013). As explained in Section 2, it is possible to interpret automation as a decline of the price of a specific equipment from infinity (the machine does not exist) to 1. Yet, our assumption that once a machine is invented, its price is constant, is crucial for deriving the general conditions under which the real wages of low-skill workers must increase asymptotically in Proposition 2. We generalize this in what follows.

Let there be two final good sectors, both perfectly competitive employing CES production technology with identical elasticity of substitution, σ. The output of sector 1,

Y, is used for consumption. The output of sector 2, X, is used for machines. The two final good sectors use distinct versions of the same set of intermediate products, where we denote the use of products as $y_1(i)$ and $y_2(i)$, respectively, with $i \in [0, N]$. The two versions of product i are produced by the same supplier using production technologies that differ only in the weight on high-skill labor:

$$y_k(i) = \left[l_k(i)^{\frac{\epsilon-1}{\epsilon}} + \alpha(i)(\tilde{\varphi} x_k(i))^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon\beta_k}{\epsilon-1}} h_k(i)^{1-\beta_k},$$

where a subscript, $k = 1, 2$, refers to the sector where the product is used. Importantly, we assume $\beta_2 \geq \beta_1$, such that the production of machines relies more heavily on machines as inputs than the production of the consumption good. Continuing to normalize the price of final good Y to 1, such that the real price of machines is p_t^x, and allowing for the natural extensions of market clearing conditions, we derive below the following generalization of Proposition 2 (where $\psi_k = (\sigma - 1)^{-1}(1 - \beta_k)^{-1}$).

Proposition 6. *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:*

$$g_\infty^{p^x} = -\psi_2 (\beta_2 - \beta_1) g_\infty^N$$

$$g_\infty^{GDP} = \left[\psi_1 + \psi_1 \frac{\beta_1 (\beta_2 - \beta_1)}{1 - \beta_2} \right] g_\infty^N, \tag{81}$$

and if $G_\infty < 1$ then the asymptotic growth rate of w_{Lt} is[40]

$$g_\infty^{w_L} = \frac{1}{1 + \beta_1(\sigma - 1)} \frac{1 - \beta_2 + \beta_1 (\beta_2 - \beta_1) \left(1 - \psi_1^{-1}\right)}{1 - \beta_2 + \beta_1 (\beta_2 - \beta_1)} g_\infty^{GDP}. \tag{82}$$

Proposition 6 reduces to Proposition 2 when $\beta_2 = \beta_1$. When $\beta_2 > \beta_1$, the productivity of machine production increases faster than that of the production of Y, implying a gradual decline in the real price of machines. For given g_∞^N, a faster growth in the supply of machines increases the (positive) growth in the relative price of low-skill workers compared with machines, w_L/p^x, but simultaneously, it reduces the real price of machines, p^x. The combination of these two effects always implies that low-skill workers capture a lower fraction of the growth in Y. Low-skill wages are more likely to fall asymptotically

[40] If G_t tends towards 1 sufficiently fast such that $lim_{t\to\infty}(1 - G_t)N_t^{\psi_2(1-\mu_1)\frac{\epsilon-1}{\epsilon}}$ is finite, then $g_\infty^{w_L} = \frac{1}{\epsilon}\left(1 - \frac{(\beta_2-\beta_1)(\epsilon-1)}{(1-\beta_2+\beta_1)}\right) g_\infty^{GDP} \geq g_\infty^{p^x}$ whether ϵ is finite or not. It is clear that there always exists an ϵ sufficiently high for the real wage of low-skill workers to decline asymptotically.

for higher values of the elasticity of substitution between products, σ, as this implies a more rapid substitution away from non-automated products.

Proof. The analysis follows similar steps as in the baseline model. The cost function (3) now becomes

$$c_k\left(\alpha\left(i\right)\right) = \beta_k^{-\beta_\kappa}\left(1-\beta_k\right)^{-(1-\beta_k)}\left(w_L^{1-\epsilon}+\varphi\left(p^x\right)^{1-\epsilon}\alpha\left(i\right)\right)^{\frac{\beta_k}{1-\epsilon}} w_H^{1-\beta_k}, \tag{83}$$

for $k \in \{1,2\}$ indexing, respectively, the production of final good and machines. As before aggregating (83) and the price normalization gives a "productivity" condition, which replaces (6).

$$\left(G\left(w_L^{1-\epsilon}+\varphi\left(p^x\right)^{1-\epsilon}\right)^{\mu_1}+(1-G)w_L^{\beta_1(1-\sigma)}\right)^{\frac{1}{1-\sigma}} w_H^{1-\beta_1} = \frac{\sigma-1}{\sigma}\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}N^{\frac{1}{\sigma-1}}, \tag{84}$$

where we generalize the definition of μ: $\mu_k \equiv \frac{\beta_k(\sigma-1)}{\epsilon-1}$. Following the same methodology for the production of machines, we get

$$\left(G\left(w_L^{1-\epsilon}+\varphi\left(p^x\right)^{1-\epsilon}\right)^{\mu_2}+(1-G)w_L^{\beta_2(1-\sigma)}\right)^{\frac{1}{1-\sigma}} w_H^{1-\beta_2} = \frac{\sigma-1}{\sigma}\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}N^{\frac{1}{\sigma-1}}p^x. \tag{85}$$

Taking the ratio between these two expressions, we get

$$\frac{\left(G\left(\left(\frac{w_L}{p^x}\right)^{1-\epsilon}+\varphi\right)^{\mu_2}+(1-G)\left(\frac{w_L}{p^x}\right)^{\beta_2(1-\sigma)}\right)^{\frac{1}{1-\sigma}} w_H^{\beta_1-\beta_2}}{\left(G\left(\left(\frac{w_L}{p^x}\right)^{1-\epsilon}+\varphi\right)^{\mu_1}+(1-G)\left(\frac{w_L}{p^x}\right)^{\beta_1(1-\sigma)}\right)^{\frac{1}{1-\sigma}}} = \frac{\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}\left(p^x\right)^{1-\beta_2+\beta_1}}{\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}}. \tag{86}$$

The share of revenues accruing to machines in the production of product i for the usage-k (i.e for use in the final sector or the machines sector) is given by

$$\nu_{k,x}\left(\alpha\left(i\right)\right) = \frac{\sigma-1}{\sigma}\alpha\left(i\right)\beta_k\frac{\varphi\left(p^x\right)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi\left(p^x\right)^{1-\epsilon}}, \tag{87}$$

aggregating over all products and denoting $R_k\left(\alpha\left(i\right)\right)$ the revenues generated through usage k by a firm of type $\alpha\left(i\right)$, we get that the total expenses in machines are given by

$$p^x X = NG\left(R_1(1)\nu_{1,x}(1)+R_2(1)\nu_{2,x}(1)\right). \tag{88}$$

The zero profit condition in the machines sector gives

$$p^x X = N\left(GR_2\left(1\right) + \left(1-G\right)R_2\left(0\right)\right). \tag{89}$$

Revenues themselves are given by

$$R_1\left(\alpha\left(i\right)\right) = \left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1} c_1\left(\alpha\left(i\right)\right)^{1-\sigma} Y \text{ and } R_2\left(\alpha\left(i\right)\right) = \left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1} c_2\left(\alpha\left(i\right)\right)^{1-\sigma} p^x X, \tag{90}$$

so that (8) still holds but separately for revenues occurring from each activity and with μ_k replacing μ. Combining (8), (87), (88) and (89), we get

$$\begin{aligned}\left(G\left(1-\tfrac{\sigma-1}{\sigma}\beta_2\tfrac{\varphi(p^x)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi(p^x)^{1-\epsilon}}\right)+(1-G)\left(1+\varphi\left(\tfrac{w_L}{p^x}\right)^{\epsilon-1}\right)^{-\mu_2}\right)\tfrac{R_2(1)}{R_1(1)} \\ = G\tfrac{\sigma-1}{\sigma}\beta_1\tfrac{\varphi(p^x)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi(p^x)^{1-\epsilon}}\end{aligned}, \tag{91}$$

which determines the revenues ratio as a function of input prices solely.

To derive low-skill wages, we compute the share of revenues accruing to low-skill labor in the production of product i for the usage-k as:

$$\nu_{k,l}\left(\alpha\left(i\right)\right) = \frac{\sigma-1}{\sigma}\beta_k\left(1+\alpha\left(i\right)\varphi\left(\frac{w_L}{p^x}\right)^{\epsilon-1}\right)^{-1},$$

so that total low-skill income can be written as:

$$w_L L = N\left(GR_1(1)\nu_{1,l}(1) + (1-G)R_1(0)\nu_{1,l}(0) + GR_2(1)\nu_{2,l}(1) + (1-G)R_2(0)\nu_{2,l}(0)\right). \tag{92}$$

The share of revenues going to high-skill workers is given by $\nu_{k,h} = \frac{\sigma-1}{\sigma}\left(1-\beta_k\right)$ both in automated and non-automated firms. As a result

$$w_H H^P = N\left(\nu_{1,h}\left(GR_1(1) + (1-G)R_1(0)\right) + \nu_{2,h}\left(GR_2(1) + (1-G)R_2(0)\right)\right), \tag{93}$$

Take the ratio between (92) and (93), and use (8) to obtain:

$$\frac{w_L L}{w_H H^P} = \frac{\left\{ \begin{array}{c} \beta_1 \left(G \left(1 + \varphi \left(\frac{w_L}{p^x} \right)^{\epsilon-1} \right)^{-1} + (1-G) \left(1 + \varphi \left(\frac{w_L}{p^x} \right)^{\epsilon-1} \right)^{-\mu_1} \right) \\ + \beta_2 \frac{R_2(1)}{R_1(1)} \left(G \left(1 + \varphi \left(\frac{w_L}{p^x} \right)^{\epsilon-1} \right)^{-1} + (1-G) \left(1 + \varphi \left(\frac{w_L}{p^x} \right)^{\epsilon-1} \right)^{-\mu_2} \right) \end{array} \right\}}{\left\{ \begin{array}{c} (1-\beta_1) \left(G + (1-G) \left(1 + \varphi \left(\frac{w_L}{p^x} \right)^{\epsilon-1} \right)^{-\mu_1} \right) \\ + (1-\beta_2) \frac{R_2(1)}{R_1(1)} \left(G + (1-G) \left(1 + \varphi \left(\frac{w_L}{p^x} \right)^{\epsilon-1} \right)^{-\mu_2} \right) \end{array} \right\}}. \tag{94}$$

Together (84), (86), (91) and (94) determine w_L, w_H, p^x and $R(2)/R(1)$ given N, G and H^P.

Asymptotic behavior for $\epsilon < 1$. As the supply of machines is going up and there is imperfect substitutability in production between machines and low-skill labor, any equilibrium must feature $w_{L\infty}/p^x_\infty = \infty$ even if $w_{L\infty} < \infty$. Applying this to (86), we get

$$(p^x_t)^{1-\beta_2+\beta_1} \sim \frac{\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}}{\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}} \varphi^{\frac{\mu_2-\mu_1}{1-\sigma}} w_{Ht}^{\beta_1-\beta_2}. \tag{95}$$

Further plugging this last relationship in (84), we get:

$$w_{Ht} \sim \frac{\left(\frac{\sigma-1}{\sigma}\right)^{1+\frac{\beta_1}{1-\beta_2}} \varphi^{\psi_2\mu_1} \beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1} \left(\beta_2^{\frac{\beta_2}{1-\beta_2}}(1-\beta_2) \right)^{\beta_1}}{G_t^{\psi_1\left(1+\frac{\beta_1(\beta_2-\beta_1)}{(1-\beta_2)}\right)} N_t^{\psi_1\left(1+\frac{\beta_1(\beta_2-\beta_1)}{(1-\beta_2)}\right)}}, \tag{96}$$

Hence

$$g^{w_H}_\infty = \psi_1 \left(1 + \frac{\beta_1(\beta_2-\beta_1)}{(1-\beta_2)} \right) g^N_\infty. \tag{97}$$

Through (91), the revenues of the machines sector and the final good sector are of the same order, which implies that Y, $p^x X$ and w_H grow at the same rate. Therefore

$$g^{GDP}_\infty = g^Y_\infty = g^{w_H}_\infty = \psi_1 \left(1 + \frac{\beta_1(\beta_2-\beta_1)}{(1-\beta_2)} \right) g^N_\infty.$$

In fact (91) gives

$$\frac{R_{2,t}(1)}{R_{1,t}(1)} \sim \frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}. \tag{98}$$

Using (95) and (96), one further gets:

$$p_t^x \sim \frac{\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}}{\left(\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}\right)^{\frac{1-\beta_1}{1-\beta_2}}}\varphi^{\psi_2\mu_1\frac{(\beta_1-\beta_2)}{\beta_1}}\left(\frac{\sigma}{\sigma-1}\right)^{\frac{\beta_2-\beta_1}{1-\beta_2}} G_t^{-\psi_2(\beta_2-\beta_1)}N_t^{-\psi_2(\beta_2-\beta_1)},$$

therefore

$$g_\infty^{p_x} = -\psi_2\left(\beta_2-\beta_1\right)g_\infty^N < 0, \tag{99}$$

since $\beta_2 > \beta_1$. Using that $w_{L\infty}/p_\infty^x = \infty$ and (98) in (94) leads to:

$$w_{Lt}\left(\frac{w_{Lt}}{p_t^x}\right)^{\epsilon-1} \sim \frac{w_{Ht}H_t^P\left(\begin{array}{c}\beta_1\left(G_t+(1-G_t)\left(\varphi\left(\frac{w_{Lt}}{p_t^x}\right)^{\epsilon-1}\right)^{1-\mu_1}\right)\\ +\beta_2\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\left(G_t+(1-G_t)\left(\varphi\left(\frac{w_{Lt}}{p_t^x}\right)^{\epsilon-1}\right)^{1-\mu_2}\right)\end{array}\right)}{\varphi G_t L\left(1-\beta_1+(1-\beta_2)\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\right)}. \tag{100}$$

Since $\beta_2 > \beta_1$, then $(1-G_t)\left(\frac{w_{Lt}}{p_t^x}\right)^{(\epsilon-1)(1-\mu_1)}$ dominates $(1-G_t)\left(\frac{w_{Lt}}{p_t^x}\right)^{(\epsilon-1)(1-\mu_2)}$ asymptotically regardless of the value of G_∞ (in other words, we can always ignore $(1-G_t)\left(\frac{w_{Lt}}{p_t^x}\right)^{(\epsilon-1)(1-\mu_2)}$ in our analysis).

The reasoning then follows that of Appendix 6.2. If $G_\infty < 1$, then (100) implies

$$w_{Lt}^{1+\beta_1(\sigma-1)} \sim \left(p_t^x\right)^{(\sigma-1)\beta_1}\frac{w_{Ht}H^P\beta_1\left(1-G_t\right)}{\varphi^{\mu_1}G_tL\left(1-\beta_1+(1-\beta_2)\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\right)}, \tag{101}$$

which, together with (97) and (99) gives (82).

Alternatively assume that $G_\infty = 1$ and that $\lim\left(1-G_t\right)N_t^{\psi_2(1-\mu_1)\frac{\epsilon-1}{\epsilon}}$ exists and is finite. Suppose first that $\limsup\left(1-G_t\right)\left(\frac{w_{Lt}}{p_t^x}\right)^{(\epsilon-1)(1-\mu_1)} = \infty$, then there must be a sub-sequence where (101) is satisfied, which with (97) and (99) leads to a contradiction with the assumption that $\lim\left(1-G_t\right)N_t^{\psi_2(1-\mu)\frac{\epsilon-1}{\epsilon}}$ exists and is finite.

If $\lim\left(1-G\right)\left(\frac{w_{Lt}}{p^x}\right)^{(\epsilon-1)(1-\mu_1)} = 0$, then (100) gives

$$w_{Lt}^{\epsilon} \sim \frac{\left(p_t^x\right)^{\epsilon-1}w_{Ht}H_t^P\left(\beta_1+\beta_2\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\right)}{\varphi L\left(1-\beta_1+(1-\beta_2)\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\right)},$$

which implies with (97) and (99) that:

$$g_{\infty}^{w_L} = \frac{1}{\epsilon}\left(1 - \frac{(\beta_2 - \beta_1)(\epsilon - 1)}{(1 - \beta_2 + \beta_1)}\right) g_{\infty}^{GDP}. \tag{102}$$

Finally, if $\limsup (1 - G_t)\, w_{Lt}^{(\epsilon-1)(1-\mu)}$ is finite but strictly positive, then as in Appendix 6.2, one can show that this requires that $\lim (1 - G_t)\, N_t^{\frac{\psi_2}{\epsilon}(\epsilon-1)(1-\mu_1)} > 0$, from which we can derive that (102) also holds in that case. This proves Proposition 6 and the associated footnote in the imperfect substitutes case.

Perfect substitutes case. In the perfect substitutes case, (84) becomes:

$$\begin{aligned}&\left(G\widetilde{\varphi}^{\beta_1(\sigma-1)}(p^x)^{\beta_1(1-\sigma)} + (1-G)w_L^{\beta_1(1-\sigma)}\right)^{\frac{1}{1-\sigma}} w_H^{1-\beta_1} \\ &= \tfrac{\sigma-1}{\sigma}\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}N^{\frac{1}{\sigma-1}} \text{ for } w_L > p^x/\widetilde{\varphi}\end{aligned}, \tag{103}$$

$$w_L^{\beta_1} w_H^{1-\beta_1} = \frac{\sigma - 1}{\sigma}\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}N^{\frac{1}{\sigma-1}} \text{ for } w_L < p^x; \tag{104}$$

(86) becomes

$$\frac{\left(G\widetilde{\varphi}^{\beta_2(\sigma-1)}(p^x)^{\beta_2(1-\sigma)} + (1-G)w_L^{\beta_2(1-\sigma)}\right)^{\frac{1}{1-\sigma}} w_H^{\beta_1-\beta_2}}{\left(G\widetilde{\varphi}^{\beta_1(\sigma-1)}(p^x)^{\beta_1(1-\sigma)} + (1-G)w_L^{\beta_1(1-\sigma)}\right)^{\frac{1}{1-\sigma}}} = \frac{\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}p^x}{\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}} \text{ for } w_L > p^x/\widetilde{\varphi}, \tag{105}$$

$$p_x = \frac{\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}}{\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}} w_L^{\beta_2-\beta_1} w_H^{\beta_1-\beta_2} \text{ for } w_L < p^x/\widetilde{\varphi}; \tag{106}$$

(91) becomes

$$\left(G\left(1 - \frac{\sigma-1}{\sigma}\beta_2\right) + (1-G)\,\widetilde{\varphi}^{\beta_2(1-\sigma)}\left(\frac{w_L}{p^x}\right)^{\beta_2(1-\sigma)}\right)\frac{R_2(1)}{R_1(1)} = G\frac{\sigma-1}{\sigma}\beta_1 \text{ for } w_L > p^x/\widetilde{\varphi}, \tag{107}$$

with $R_2(1) = 0$ for $w_L < p^x/\widetilde{\varphi}$; and (94) becomes

$$\frac{w_L L}{w_H H^P} = (1-G)\frac{\beta_1\left(\widetilde{\varphi}\frac{w_L}{p^x}\right)^{\beta_1(1-\sigma)} + \beta_2\frac{R_2(1)}{R_1(1)}\left(\widetilde{\varphi}\frac{w_L}{p^x}\right)^{\beta_2(1-\sigma)}}{\left\{\begin{array}{c}(1-\beta_1)\left(G + (1-G)\left(\widetilde{\varphi}\frac{w_L}{p^x}\right)^{\beta_1(1-\sigma)}\right) \\ +(1-\beta_2)\frac{R_2(1)}{R_1(1)}\left(G + (1-G)\left(\widetilde{\varphi}\frac{w_L}{p^x}\right)^{\beta_2(1-\sigma)}\right)\end{array}\right\}} \text{ for } w_L > p^x/\widetilde{\varphi}, \tag{108}$$

$$\frac{w_L L}{w_H H^P} = \frac{\beta_1}{1-\beta_1} \text{ for } w_L < p^x/\widetilde{\varphi}. \tag{109}$$

Together (104), (106) and (109) show that we must have $w_{Lt} \geq \frac{p_t^x}{\widetilde{\varphi}}$ for t large enough, which delivers (97) and (99).

Assume that $G_\infty < 1$, then (108) gives (101) from which we get that (82) is satisfied.

Now consider the case where $G_\infty = 1$ and $\lim (1-G_t) N_t^{\psi_2}$ exists and is finite. Then (108) and (107) imply

$$w_{Lt} \sim (1-G_t)\, w_{Ht} \left(\widetilde{\varphi}\frac{w_{Lt}}{p_t^x}\right)^{\beta_1(1-\sigma)} \frac{\left(\beta_1 + \beta_2 \frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\left(\widetilde{\varphi}\frac{w_{Lt}}{p_t^x}\right)^{-(\beta_2-\beta_1)(\sigma-1)}\right)}{1-\beta_1+(1-\beta_2)\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}} \frac{H_t^P}{L} \text{for } w_L > p^x/\widetilde{\varphi}.$$

We can then derive that $\frac{\widetilde{\varphi} w_{Lt}}{p_t^x}$ must have a finite (and positive) limit, so that

$$g_\infty^{w_L} = g_\infty^{p^x} = -\frac{\beta_2-\beta_1}{1-\beta_2+\beta_1} g_\infty^{GDP}.$$

This proves Proposition 6 and its associated footnote in the perfect substitutes case. □

7.5 Intermediate steps and additional results on the baseline dynamic model

We provide intermediate steps for the proofs of Section 6.3, the proof of Corollary 1 and additional analytical results on the dynamic model of Section 3.

7.5.1 Intermediate steps for section 6.3.1

In this section we derive (37). Taking the difference between (32) and (33) and using (34) we obtain:

$$\left(r_t - (\psi-1)\, g_t^N\right)\left(\widehat{V}_t^A - \widehat{V}_t^N\right) = \hat{\pi}_t^A - \hat{\pi}_t^N - \frac{1-\kappa}{\kappa}\widehat{v}_t \widehat{h}_t^A + \left(\dot{\widehat{V}}_t^A - \dot{\widehat{V}}_t^N\right).$$

Using again (34) we get,

$$r_t - (\psi - 1)\, g_t^N = \kappa\eta G_t^{\widetilde{\kappa}} \left(\widehat{h}_t^A\right)^{\kappa-1} \left[\frac{\hat{\pi}_t^A - \hat{\pi}_t^N}{\widehat{v}_t} - \frac{1-\kappa}{\kappa}\widehat{h}_t^A + \frac{d}{dt}\left(\frac{\left(\widehat{h}_t^A\right)^{1-\kappa}}{\kappa\eta G_t^{\widetilde{\kappa}}}\right)\right] + \frac{\dot{\widehat{v}}_t}{\widehat{v}_t}.$$

Using (35), we can rewrite this expression as

$$\gamma\left(\frac{\widehat{\pi}_t^N}{\widehat{v}_t} + \frac{1-\kappa}{\kappa}\widehat{h}_t^A\right) = \kappa\eta G_t^{\widetilde{\kappa}} \left(\widehat{h}_t^A\right)^{\kappa-1} \left[\frac{\hat{\pi}_t^A - \hat{\pi}_t^N}{\widehat{v}_t} - \frac{1-\kappa}{\kappa}\widehat{h}_t^A + \frac{d}{dt}\left(\frac{\left(\widehat{h}_t^A\right)^{1-\kappa}}{\kappa\eta G_t^{\widetilde{\kappa}}}\right)\right]$$

Using (31), this leads to:

$$\gamma\left(\frac{\widehat{\pi}_t^N}{\widehat{v}_t} + \frac{1-\kappa}{\kappa}\widehat{h}_t^A\right) = \begin{array}{c} \kappa\eta G_t^{\widetilde{\kappa}} \left(\widehat{h}_t^A\right)^{\kappa-1} \left[\frac{\hat{\pi}_t^A - \hat{\pi}_t^N}{\widehat{v}_t} - \frac{1-\kappa}{\kappa}\widehat{h}_t^A\right] \\ + (1-\kappa)\frac{\dot{\widehat{h}}_t^A}{\widehat{h}_t^A} - \frac{\widetilde{\kappa}}{G_t}\left(\eta G_t^{\widetilde{\kappa}} \left(\hat{h}_t^A\right)^{\kappa} (1 - G_t) - G_t g_t^N\right) \end{array}.$$

Reordering terms and using (36) gives (37).

7.5.2 Uniqueness of the steady state

Generally the steady state is not unique. Nonetheless, consider the special case in which $\widetilde{\kappa} = 0$. Then f can be rewritten as

$$f\left(g^{N*}\right) = \frac{1-\kappa}{\kappa}\frac{\gamma G^* \widehat{h}^{A*}}{\psi H^{P*}}\left(\frac{1}{\kappa\eta}\left(\widehat{h}^{A*}\right)^{1-\kappa} + \frac{1}{\gamma}\right), \tag{110}$$

note that H^{P*} is decreasing in g^{N*} and $\widehat{h}^{A*}$ is increasing in g^{N*}, so a sufficient condition for f to be increasing in g^{N*} is that $G^*\widehat{h}^{A*}$ is also increasing in g^{N*}. With $\widetilde{\kappa} = 0$, using (50), (49), we get:

$$G^*\widehat{h}^{A*} = \frac{\eta\left(\frac{\kappa}{\gamma(1-\kappa)}\right)^{\kappa+1}\left(\rho + ((\theta-1)\,\psi + 1)\, g^{N*}\right)^{\kappa+1}}{\eta\left(\frac{\kappa}{\gamma(1-\kappa)}\right)^{\kappa}\left(\rho + ((\theta-1)\,\psi + 1)\, g^{N*}\right)^{\kappa} + g^{N*}}.$$

Therefore

$$\frac{d\left(G^{*}\widehat{h}^{A*}\right)}{dg^{N*}}=\begin{array}{c}\frac{\eta\left(\frac{\kappa}{\gamma(1-\kappa)}\right)^{\kappa+1}\left(\rho+((\theta-1)\psi+1)g^{N*}\right)^{\kappa}}{\left(\eta\left(\frac{\kappa}{\gamma(1-\kappa)}\right)^{\kappa}(\rho+((\theta-1)\psi+1)g^{N*})^{\kappa}+g^{N*}\right)^{2}}\\ \times\left(\begin{array}{c}\eta\left(\frac{\kappa}{\gamma(1-\kappa)}\right)^{\kappa}((\theta-1)\psi+1)\left(\rho+((\theta-1)\psi+1)g^{N*}\right)^{\kappa}\\ -\rho+g^{N*}\kappa\left((\theta-1)\psi+1\right)\end{array}\right)\end{array}.$$

Since $g^{N*}>0$, we get that $\frac{d\left(G^{*}\widehat{h}^{A*}\right)}{dg^{N*}}>0$ (so that the steady state is unique) if $\frac{(1-\kappa)^{\kappa}\gamma^{\kappa}}{\eta\kappa^{\kappa}}\rho^{1-\kappa}<(\theta-1)\psi+1$. This condition is likely to be met for reasonable parameter values as long as the automation technology is not too concave: ρ is a small number, $\theta\geq 1$ and γ and η being innovation productivity parameters should be of the same order (it is indeed met for our baseline parameters).

7.5.3 Proof of Lemmas of section 6.3.4

Proof of Lemma 2. If $\widetilde{\kappa}=0$, $\widehat{h}_t^A$ cannot remain small forever as with positive growth in N_t, N_t and therefore w_{Lt} will become large. Since, the Poisson rate is $\eta\left(\widehat{h}_t^A\right)^{\kappa}=O\left(\varphi w_{Lt}^{\epsilon-1}\right)$. This implies that G_t must start growing at a positive rate and cannot converge toward 0.

When $\widetilde{\kappa}>0$ (and $G_0\neq 0$, otherwise automation is impossible), however, whether the Poisson rate of automation may remain negligible or not depends on a horse race between the drop in the share of automated products (and therefore the efficiency of the automation technology) and the rise in the low-skill wages (which, through horizontal innovation can become arbitrarily large).

First assume that $G_t w_{Lt}^{\beta(\sigma-1)}$ does not tend towards 0. Then from (55) we obtain that:

$$\widehat{h}_t^A=O\left(G_t^{\widetilde{\kappa}-1}\right)^{\frac{1}{1-\kappa}}\Longrightarrow\eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa}=O\left(G_t^{\frac{\widetilde{\kappa}-\kappa}{1-\kappa}}\right)\tag{111}$$

Since $\widetilde{\kappa}\leq\kappa$, we obtain that the Poisson rate of automation increases without bound, so G_t cannot converge toward 0.

Assume now that $G_t w_{Lt}^{\beta(\sigma-1)}$ does tend towards 0. This ensures that $w_{Lt}=O\left(N_t^{\frac{1}{\sigma-1}}\right)$. Moreover, $\frac{\pi_t^A-\pi_t^N}{G_t\pi_t^A+(1-G_t)\pi_t^N}=O\left(w_{Lt}^{\beta(\sigma-1)}\right)$. Then using this in (55), we obtain

$$\widehat{h}_t^A=O\left(G_t^{\widetilde{\kappa}}w_{Lt}^{\beta(\sigma-1)}\right)^{\frac{1}{1-\kappa}}.$$

Note that $\widehat{h}_t^A$ must remain bounded otherwise high-skill labor market clearing is violated. Therefore, we must have $G_t^{\widetilde{\kappa}} w_{Lt}^{\beta(\sigma-1)}$ bounded (which implies that $G_t w_{Lt}^{\beta(\sigma-1)}$ tends towards 0). Therefore the Poisson rate obeys:

$$\eta G_t^{\widetilde{\kappa}} \left(\widehat{h}_t^A\right)^{\kappa} = O\left(G_t^{\frac{\widetilde{\kappa}}{1-\kappa}} N_t^{\frac{\beta\kappa}{1-\kappa}}\right)$$

Plugging this in (31) we get:

$$\dot{G}_t = O\left(G_t^{\frac{\widetilde{\kappa}}{1-\kappa}} N_t^{\frac{\beta\kappa}{1-\kappa}}\right) - g_t^N G_t$$

To obtain that the share G_t is going towards 0, it must first be that $G_t^{\frac{\widetilde{\kappa}}{1-\kappa}} N_t^{\frac{\beta\kappa}{1-\kappa}}$ declines at the same rate or faster than G_t.

Consider first the case where, $G_t^{\frac{\widetilde{\kappa}}{1-\kappa}} N_t^{\frac{\beta\kappa}{1-\kappa}}$ and G_t are of the same order. In that case, we must have:

$$G_t = O\left(N_t^{\frac{\beta\kappa}{1-\kappa-\widetilde{\kappa}}}\right)$$

This cannot go towards 0 if $1-\kappa-\widetilde{\kappa} > 0$. In addition, recall that this reasoning assumed that $G_t^{\widetilde{\kappa}} w_{Lt}^{\beta(\sigma-1)}$ remains bounded. We have

$$G_t^{\widetilde{\kappa}} w_{Lt}^{\beta(\sigma-1)} = O\left(N_t^{\frac{\beta(1-\kappa)(1-\widetilde{\kappa})}{1-\kappa-\widetilde{\kappa}}}\right),$$

which is indeed declining if $1-\kappa-\widetilde{\kappa} < 0$. Furthermore, in that case we must have $\dot{G}_t \geq -g_t^N G_t$, that is G_t should not decline at a rate faster than N_t^{-1}. This implies that we must have $\frac{\beta\kappa}{\kappa+\widetilde{\kappa}-1} \leq 1 \Longleftrightarrow \kappa(1-\beta)+\widetilde{\kappa} \geq 1$.

Alternatively, if $G_t^{\frac{\widetilde{\kappa}}{1-\kappa}} N_t^{\frac{\beta\kappa}{1-\kappa}}$ goes towards 0 faster than G_t then G_t will be declining at the rate g_t^N, so that we have $G_t = O\left(N_t^{-1}\right)$. This then implies

$$G_t^{\frac{\widetilde{\kappa}}{1-\kappa}} N_t^{\frac{\beta\kappa}{1-\kappa}} / G_t = O\left(N_t^{\frac{\beta\kappa+1-\kappa-\widetilde{\kappa}}{1-\kappa}}\right).$$

As soon as $\kappa(1-\beta)+\widetilde{\kappa} < 1$ then this cannot go towards 0.

Therefore $\kappa(1-\beta)+\widetilde{\kappa} < 1$ is a sufficient condition which ensures that the Poisson rate of automation must take off.

Proof of Lemma 3. Assume that G_t is bounded above 0 (note that Lemma **2** shows that as long as $\kappa(1-\beta)+\widetilde{\kappa} < 1$, it is impossible to have $G_\infty = 0$). Note that H_t^P

must be bounded below otherwise there would be arbitrarily large welfare gains from increasing consumption at time t and reducing it at later time periods. As H_t^P is also bounded above (by H), then we must have (following the reasoning of Appendix 6.2), that $w_{Ht} = \Theta\left(N_t^{\psi}\right)$, $C_t = \Theta\left(N_t^{\psi}\right)$ and w_{Lt} is bounded below, so that $\widehat{v}_t$ and $\widehat{c}_t$ are bounded above and below and ω_t must be bounded above.

Integrating (14), using the transversality condition and dividing by w_{Ht}/N_t, we get:

$$\frac{V_t^A}{w_{Ht}/N_t} = \int_t^{\infty} \exp\left(-\int_t^s r(u)\,du\right) \frac{\pi_s^A}{w_{Ht}/N_t} ds,$$

using the Euler equation (18), this leads to:

$$\frac{V_t^A}{w_{Ht}/N_t} = \int_t^{\infty} \exp\left(-\rho\left(t-s\right)\right) \left(\frac{C\left(s\right)}{C\left(t\right)}\right)^{-\theta} \frac{\widehat{\pi}_s^A N_s^{\psi-1}}{\widehat{v}_s \widehat{N}_t^{\psi-1}} ds.$$

Rewriting this expression with the normalized variables and using (43), we get:

$$\frac{V_t^A}{w_{Ht}/N_t} = \int_t^{\infty} e^{-\rho(s-t)} \left(\frac{N_s}{N_t}\right)^{-(1+(\theta-1)\psi)} \frac{\widehat{c}\left(s\right)^{\theta}}{\widehat{c}\left(t\right)^{\theta}} \frac{\psi\left(\varphi + \left(\omega_s n_s\right)^{\frac{1}{\mu}}\right)^{\mu} H_s^P}{G_s\left(\varphi_s + \left(\omega_s n_s\right)^{\frac{1}{\mu}}\right)^{\mu} + \left(1-G_s\right)\omega_s n_s} \frac{\widehat{v}_s}{\widehat{v}_t} ds.$$

Note that $\frac{\widehat{c}(s)}{\widehat{c}(t)}$, $\frac{\psi\left(\varphi+(\omega_s n_s)^{\frac{1}{\mu}}\right)^{\mu} H_s^P}{G\left(\varphi_s+(\omega_s n_s)^{\frac{1}{\mu}}\right)^{\mu}+(1-G_s)\omega_s n_s}$ and $\frac{\widehat{v}_s}{\widehat{v}_t}$ are all bounded and that N_s is weakly increasing, therefore we get that

$$\frac{V_t^A}{w_{Ht}/N_t} \leq \int_t^{\infty} e^{-\rho(s-t)} M ds,$$

for some constant M. This ensures that $\frac{V_t^A - V_t^N}{w_{Ht}/N_t}$ must remain bounded, and following (16), $\hat{h}_t^A$ must be bounded as well.

7.5.4 Behavior close to the steady-state

We now provide details on the behavior of the economy close to the steady-state. In the steady-state, using (50) we get

$$g^{N*} = \frac{1}{(\theta-1)\psi+1}\left(\frac{\gamma\left(1-\kappa\right)}{\kappa}\widehat{h}^{A*} - \rho\right).$$

Therefore close to the steady-state, we obtain that N_t grows at rate $g_t^N = g^{N*} + o(1)$, that the share of automated product obeys $G_t = G^* + o(1)$, with the mass of high-skill workers in automation given by $H_t^A = (1 - G^*)\widehat{h}^{A*} + o(1)$ and the mass of high-skill workers in production given by $H_t^P = H^{P*} + o(1)$, with H^{P*} given by (51). Using (41), we obtain that $\hat{v}$ is a constant in steady-state and that wages close to the steady-state obey:

$$w_{Ht} = (1-\beta)\left(\frac{\sigma-1}{\sigma}\beta^{\beta}\right)^{\frac{1}{1-\beta}} (G^*\varphi^{\mu})^{\psi} N_t^{\psi} + o\left(N_t^{\psi}\right),$$

$$w_{Lt} = (\omega^*)^{\frac{1}{\beta(1-\sigma)}} N_t^{\frac{\psi}{1+\beta(\sigma-1)}} + o\left(N_t^{\frac{\psi}{1+\beta(\sigma-1)}}\right).$$

Using (63), $\hat{\pi}^A$ is constant in steady-state and the profits made by an automated firm close to the steady-state are given by

$$\pi_t^A = \frac{1}{\sigma}\left(\frac{(\sigma-1)\beta\widetilde{\varphi}^{\frac{\epsilon}{\epsilon-1}}}{\sigma}\right)^{\frac{\beta}{1-\beta}} H^{P*}(G^*)^{\psi-1} N_t^{\psi-1} + o\left(N_t^{\psi-1}\right).$$

(36) then implies that the profits made by a non-automated firm π_t^N are negligible in front of π_t^A, with

$$\pi_t^N = w_{Lt}^{\beta(1-\sigma)}\varphi^{-\mu}\pi_t^A + o\left(w_{Lt}^{\beta(1-\sigma)}\pi_t^A\right).$$

Therefore, we get $g_\infty^{\pi^N} = g_\infty^{\pi^A} - \beta(\sigma-1) g_\infty^{w_L}$. Using (32), the value of an automated firm is then simply given by:

$$V_t^A = \frac{\pi_t^A}{r^* - (\psi-1) g^{N*}} + o\left(N_t^{\psi-1}\right), \tag{112}$$

$$\text{where } r^* = \rho + \theta\psi g^{N*} \tag{113}$$

is the steady-state interest rate, so that $g_\infty^{V^A} = g_\infty^{\pi^A}$. Following (33) and (34), the normalized value of a non-automated firm obeys:

$$\left(r_t - (\psi-1) g_t^N\right)\widehat{V}_t^N = \widehat{\pi}_t^N + (1-\kappa)\eta G_t^{\widetilde{\kappa}}\widehat{h}_t^A\left(\widehat{V}_t^A - \widehat{V}_t^N\right) + \dot{\widehat{V}}_t^N.$$

Therefore, one gets that for large N_t,

$$V_t^N = \frac{(1-\kappa)\,\eta G^{*\widetilde{\kappa}}\widehat{h}^{A*}}{r^* - (\psi-1)\,g^{N*} + (1-\kappa)\,\eta G^{*\widetilde{\kappa}}\widehat{h}^{A*}} V_t^A + o\left(N_t^{\psi-1}\right), \tag{114}$$

so that asymptotically all the value of a new firm comes from the profits it makes once automated and $g_\infty^{V^N} = g_\infty^{V^A}$.

7.5.5 Proof of Corollary 1.

Appendix 7.5.4 shows that in steady-state $\widehat{V}^{N*} = f\widehat{V}^{A*}$ with

$$f = \frac{(1-\kappa)\,\eta G^{*\widetilde{\kappa}}\widehat{h}^{A*}}{\rho + (\psi\,(\theta-1)+1)\,g_N^* + (1-\kappa)\,\eta G^{*\widetilde{\kappa}}\widehat{h}^{A*}}, \tag{115}$$

$$\widehat{V}^{A*} = \frac{\widehat{\pi}_t^A}{\rho + (\psi\,(\theta-1)+1)\,g_N^*}.$$

Using $\widehat{V}^{N*} = \widehat{v}^*/\gamma$ and (63), we obtain $f\frac{1}{\rho+(\psi(\theta-1)+1)g_N^*}\frac{\psi H^{P*}}{G^*} = \frac{1}{\gamma}$. Rearranging terms and using (51), this leads to

$$g^{N*} = \frac{\frac{f}{G^*}\gamma\psi\left(H - (1-G^*)\,\widehat{h}^{A*}\right) - \rho}{\frac{f}{G^*}\psi + \frac{\theta-1}{(\sigma-1)(1-\beta)} + 1}, \tag{116}$$

while from (61) the growth rate when N_t is low is approximately given by

$$g^{N1} = \frac{\gamma H\psi - \rho}{\psi + \frac{\theta-1}{\sigma-1} + 1}. \tag{117}$$

The two expressions differ by three terms: In the numerator, $H - (1-G^*)\,\widehat{h}^{A*}$ in (116) replaces H in (117), as some high-skill workers are hired to automate close to the steady-state, the pool of high-skill workers available for horizontal innovation or production is smaller, and this force pushes toward $g^{N*} < g^{N1}$. In the denominator, $\frac{\theta-1}{(\sigma-1)(1-\beta)}$ in (116) replaces $\frac{\theta-1}{\sigma-1}$ in (117) because the growth rate in the number of products has a larger impact on the economy growth rate with automation than without. This increases the effective interest rate and reduces the present value of an automated firm, therefore it also pushes toward $g^{N*} < g^{N1}$. Finally the term f/G^* in (116) does not exist in (117). Note that $\partial g^{N*}/\left(\partial f/G^*\right) > 0$, so that this term reflects two different forces. On one hand close to the steady-state, the value of a new firm is a fraction $f < 1$ of the value of

an automated firm. On the other hand, the profits of automated firms are larger by a factor $1/G^*$ than aggregate profits, which remain a fraction $1/\sigma$of total output through the entire transitional dynamics, and this increases the value of non-automated firms. Combining (31), (49) and (115), we get

$$f/G^* < 1 \Leftrightarrow (1-\kappa)\, g^{N*} < \rho + (\psi\,(\theta-1)+1)\, g_N^*.$$

Since $\theta \geq 1$ and $\kappa < 1$, this inequality necessarily holds and $f/G^* < 1$. Interpreting the value of a new firm as the discounted flow of a "net profit flow"$f\widehat{\pi}_t^A$, the profit flow of new firms in the asymptotic steady-state is a lower fraction of total output than it is for low N_t ensuring that $g^{N*} > g^{N1}$. This establishes Corollary 1.

7.5.6 Comparative statics

In this section we establish comparative static results in the steady-state.

Proposition 7. *The asymptotic growth rates of GDP g_∞^{GDP} and low-skill wages $g_\infty^{w_L}$ increase in the productivity of automation η and horizontal innovation γ.*

Therefore, in the long-run, a better automation technology (a higher η) actually benefits low-skill wages: the reason is that firms automate faster which encourages horizontal innovation.[41] During the transition, however, a higher η also means that automation takes off sooner, leading to lower low-skill wages at that point and a higher skill premium (see a numerical example in Appendix 7.6.3). These result preview those on the effect of taxes on automation in Section 7.11.6.

Proof. The proposition is established when the steady state is unique but it extends to the case of the steady states with the highest and lowest growth rates when there is multiplicity. Recall that the steady state is characterized as the solution to an equation $f\left(g^{N*}\right) = 1$ through (53), where G^*, $\widehat{h}^{A*}$ and H^{P*} can all be written as functions of g^{N*} and parameters. Moreover, when there is a single steady state (as well as for the steady states with the highest and the lowest growth rates in case of multiplicity), f must be increasing in the neighborhood of g^{N*}.

[41] Corollary 1 establishes that the growth rate of the number of products is higher in a world with no automation at all than in a world with automation, but Proposition 7 shows that conditional on automation happening ($\eta > 0$), the asymptotic growth rate in the number of products is higher when automation is easier (η is higher).

Comparative static with respect to γ. (50) implies that $\widehat{h}^{A*}$ is inversely proportional to γ (for given g^{N*}). Formally, we have:

$$\frac{\partial \widehat{h}^{A*}}{\partial \gamma} = -\frac{\widehat{h}^{A*}}{\gamma}. \tag{118}$$

Differentiating (49) and using (118) leads to:

$$\frac{\partial G^*}{\partial \gamma} = \frac{-\kappa g^{N*} G^*}{\gamma \left(\eta \left(G^*\right)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + \left(1-\widetilde{\kappa}\right) g^{N*} \right)}, \tag{119}$$

so that for a given g^{N*}, G^* is also decreasing in γ. Using (51), (118) and (119), we get:

$$\frac{\partial H^{P*}}{\partial \gamma} = \frac{1}{\gamma} \left(\begin{array}{c} \frac{g^{N*}}{\gamma} + \left(1-G^*\right)\left(1-\kappa\right)\widehat{h}^{A*} \\ + \frac{(1-\widetilde{\kappa})\kappa \widehat{h}^{A*} G^* \left(g^{N*}\right)^2}{\left(\eta (G^*)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa}\right)\left(\eta (G^*)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + (1-\widetilde{\kappa}) g^{N*}\right)} \end{array} \right) > 0$$

so that H^{P*} is increasing in γ. Note that f, defined in (53), can be rewritten as

$$f\left(g^{N*}\right) = \frac{1-\kappa}{\kappa} \frac{1}{\psi H^{P*}} \left(\frac{\left(G^*\right)^{1-\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{1-\kappa} \left(\gamma \widehat{h}^{A*}\right)}{\kappa \eta} + G^* \widehat{h}^{A*} \right),$$

which shows that f is decreasing in γ for a given g^{N*} (H^{P*} is increasing, G^* and $\widehat{h}^{A*}$ are decreasing, and $\gamma \widehat{h}^{A*}$ is constant). Since f is increasing in g^{N*} at the equilibrium value, (53) implies that g^{N*} increases in γ.

Comparative static with respect to η. For given g^{N*}, (50) implies that $\widehat{h}^{A*}$ does not depend on η. Differentiating (49), we get:

$$\frac{\partial \ln G^*}{\partial \ln \eta} = \frac{g^{N*}}{\eta \left(G^*\right)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + \left(1-\widetilde{\kappa}\right) g^{N*}}, \tag{120}$$

so for given g^{N*}, G^* increases in η. (51) implies then that

$$\frac{\partial \ln H^{P*}}{\partial \ln \eta} = \frac{G^* \widehat{h}^{A*}}{H^{P*}} \frac{g^{N*}}{\eta \left(G^*\right)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + \left(1-\widetilde{\kappa}\right) g^{N*}}.$$

Using this equation together with (120) and (53), we obtain:

$$\frac{\partial \ln f}{\partial \ln \eta} = \left\{ \begin{array}{c} \frac{g^{N*}}{\eta (G^*)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + (1-\widetilde{\kappa}) g^{N*}} \left(1 - \frac{G^* \widehat{h}^{A*}}{H^{P*}} - \widetilde{\kappa} \frac{\frac{1}{\kappa \eta (G^*)^{\widetilde{\kappa}}} \left(\widehat{h}^{A*}\right)^{1-\kappa}}{\frac{1}{\kappa \eta (G^*)^{\widetilde{\kappa}}} \left(\widehat{h}^{A*}\right)^{1-\kappa} + \frac{1}{\gamma}} \right) \\ - \frac{\frac{1}{\kappa \eta (G^*)^{\widetilde{\kappa}}} \left(\widehat{h}^{A*}\right)^{1-\kappa}}{\frac{1}{\kappa \eta (G^*)^{\widetilde{\kappa}}} \left(\widehat{h}^{A*}\right)^{1-\kappa} + \frac{1}{\gamma}} \end{array} \right\}.$$

Using (50), we can rewrite this as:

$$\frac{\partial \ln f}{\partial \ln \eta} = \left\{ \begin{array}{c} - \frac{1}{\eta (G^*)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + (1-\widetilde{\kappa}) g^{N*}} \\ \times \left(\frac{g^{N*} G^* \widehat{h}^{A*}}{H^{P*}} + \frac{\rho + ((\theta-1)\psi + \kappa) g^{N*}}{\gamma (1-\kappa) \left(\frac{1}{\eta_L \kappa G_L^{\widetilde{\kappa}}} \left(\widehat{h}^{A*}\right)^{1-\kappa} + \frac{1}{\gamma} \right)} \right) \end{array} \right\},$$

so that f is decreasing in η. This implies that g^{N*} must be increasing in η. Since $\widehat{h}^{A*}$ only depends on η through g^{N*}, we also get that $\widehat{h}^{A*}$ increases in η. □

7.6 Numerical illustration

We illustrate the results of Section 3.3 and further analyze the behavior of our economy through the use of numerical simulations.[42] Section 7.6.2 shows examples where low-skill wages temporarily decline. Section 7.6.3 shows the effects of changing the innovation parameter on wages and Section 7.6.4 gives a systematic exploration of the parameter space. Section 4 calibrates a richer model to the U.S. data.

7.6.1 Illustrating Section 3.3

Unless otherwise specified, the broad patterns described below do not depend on specific parameter choices and we simply choose "reasonable" parameters (Table 4). For convenience we loosely refer to the time period where N_t is low and the economy behaves close to a Romer model as the first phase, the period where the economy approaches its steady-state as the third phase, and the period in between as the second phase

Baseline Parameters. Total stock of labor is 1 with $L = 2/3$ 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 set low at $N_0 = 1$ to ensure we begin in Phase

[42] We employ the so-called "relaxation" algorithm for solving systems of discretized differential equations (Trimborn, Koch and Steger, 2008). See Appendix 7.7 for details.

Table 4: Baseline Parameter Specification

σ	ϵ	β	H	L	θ	η	κ	$\tilde{\varphi}$	ρ	$\widetilde{\kappa}$	γ	N_0	G_0
3	4	2/3	1/3	2/3	2	0.2	0.5	0.25	0.02	0	0.3	1	0.001

1. The initial share of automated products is low, $G_0 = 0.001$, but would initially decline had we chosen a higher level. We set $\sigma = 3$ to capture an initial labor share close to 2/3. We set $\tilde{\varphi} = 0.25$ and $\epsilon = 4$, so that at $t = 0$, the profits of automated firms relative to non-automated firms are only 0.004%. The innovation parameters (γ, η, κ) are chosen such that GDP growth is close to 2% both initially and asymptotically. There is no externality from the share of automated products in the automation technology, $\widetilde{\kappa} = 0$. ρ and θ are chosen such that the interest rate is around 6% initially and asymptotically.

Figure 11 plots the evolution of the economy. Based on the behavior of automation expenditures (Panel C) we roughly delimit Phase 1 as corresponding to the first 100 years and Phase 2 as the period between year 100 and year 250.

Innovation and growth. Initially, low-skill wages and hence the incentive to automate—proportional to $(V_t^A - V_t^N)/(w_{Ht}/N_t)$—are low (Panel B) and so is the share of automated firms G_t (Panel C). With growing low-skill wages, the incentive to automate picks up a bit before year 100. Then the economy enters Phase 2 as automation expenses sharply increase (up to 4% of GDP). Innovation is progressively more directed toward automation (Panel C) and the share of automated products G_t rises before stabilizing at a level strictly below 1. There is no simple one-to-one link between the direction of innovation and the speed of the increase in inequality. The skill premium increases the fastest in year 180 while innovation is increasingly directed towards automation until year 192. More generally the growth rate of the skill premium declines in Phase 3 relative to the middle of Phase 2 even though the share of automation innovation stays at a high level.

In line with Proposition 1, spending on horizontal innovation as a share of GDP declines during Phase 2 and for any parameter values ends up being lower in Phase 3 than Phase 1. Despite this, the growth rate of GDP is roughly the same in Phases 1 and 3 because the lower rate of horizontal innovation in Phase 3 is compensated by a higher elasticity of GDP wrt. N_t $(1/[(\sigma-1)(1-\beta)]$ instead of $1/(\sigma-1))$. As a result, the phase of intense automation—which also contributes to growth—is associated with a temporary boost of growth. This is, however, specific to parameters.

Wages. In the first phase, growth comes mostly from horizontal innovation and both

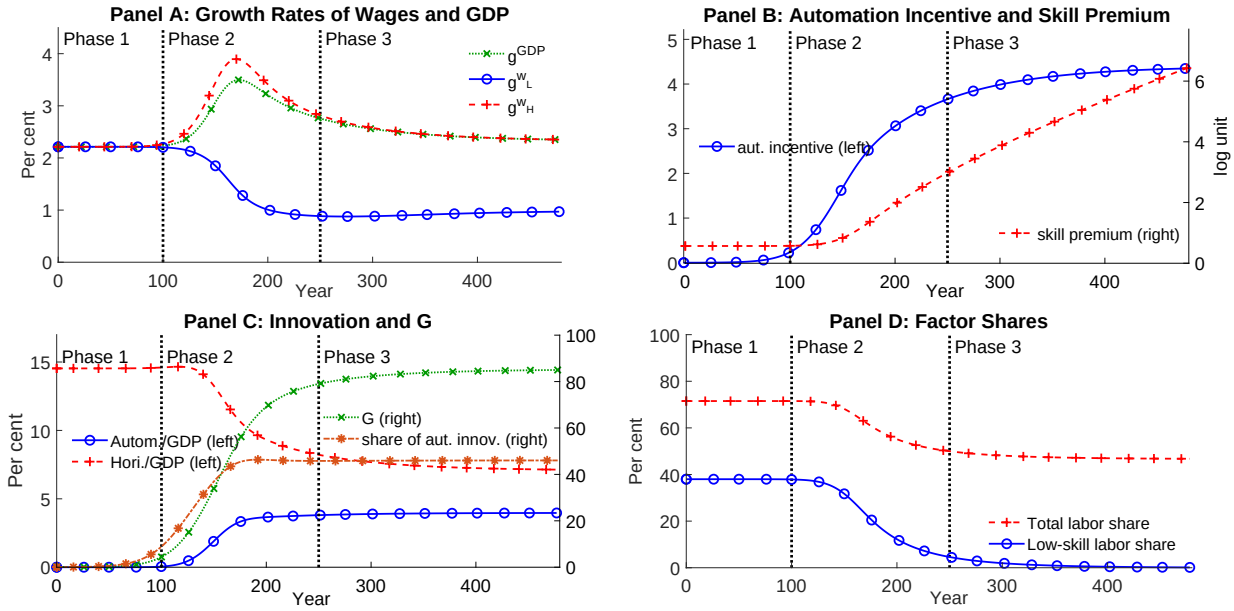

Figure 11: Transitional Dynamics for baseline parameters. Panel A shows growth rates for GDP, low-skill wages (w_L) and high-skill wages (w_H), Panel B the incentive to automate, $\left(V_t^A - V_t^N\right) / \left(w_{Ht}/N_t\right)$, and the skill premium, Panel C the total spending on horizontal innovation 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.

wages grow at around 2% (Panel A). As rising low-skill wages trigger the second phase, the growth rate of high-skill wages increases to almost 4% and the growth rate of low-skill wages declines to around 1%. Though our parameter values satisfy the conditions of Proposition 1 B.ii and any increase in G_t has a negative impact on w_{Lt}, the growth in N_t is sufficient to ensure that low-skill wages grow at a positive rate throughout (see Section 7.6.2 for counter-examples). Finally, in the third phase, the growth rate of low-skill wages stabilizes at around 1% and the skill premium keeps rising but more slowly than previously.

Factor shares. Panel D of Figure 11 plots the labor share and the low-skill labor share. With machines as intermediate inputs, capital income corresponds to aggregate profits, which are a constant share of output. High-skill labor in production also earns a constant share of output. Both correspond to a rising share of GDP in Phase 2 as during this time period, the ratio Y/GDP increases since machines expenditures are excluded from GDP. The low-skill labor share is nearly constant in Phase 1 but declines with automation in Phase 2 and approaches 0 in Phase 3. The total labor share of

GDP follows a similar pattern—but its decline is less marked since the high-skill share increases. This occurs despite an increase in the share of high-skill workers in innovation which raises the labor share (see (9)). Yet, because of this effect, the drop in the labor share can be delayed relative to the rise in the skill premium for some parameter values.

Wealth and consumption. Figure 12 shows the evolution of wealth and consumption for the baseline parameters both in the aggregate and for each skill group. Panel A shows that consumption growth follows a pattern very similar to that of GDP growth (displayed in Figure 11.A), which is in line with a stable ratio of total R&D expenses over GDP across the three phases (Figure 11.D). In the absence of any financial constraints, low-skill and high-skill consumption must grow at the same rate, with high-skill workers consuming more since they have a higher income (Panel B). Since low-skill labor income becomes a negligible share of GDP, while the high-skill labor share increases, a constant consumption ratio can only be achieved if high-skill workers borrow from low-skill workers in the long-run. This is illustrated in Panel C, which shows the share of assets held by low-skill workers, under the assumption that initially assets holdings per capita are identical for low-skill and high-skill workers (so that low-skill workers hold 2/3 of the assets in year 0, since with these parameters $H/L = 1/2$). Initially, low-skill and high-skill income grow at a constant rate so that the share of assets held by low-skill workers is stable; but, in anticipation of a lower growth rate for low-skill wages than for high-skill wages, low-skill workers start saving more, and the share of assets they hold increases. This share eventually reaches more than 100%, meaning that the high-skill workers net worth becomes negative. Panel D shows that since profits become a higher share of GDP (an effect which dominates a temporary increase in the interest rate in Phase 2), the wealth to GDP ratio increases in phase 2, such that its steady state value is nearly 3 times higher than its original value.

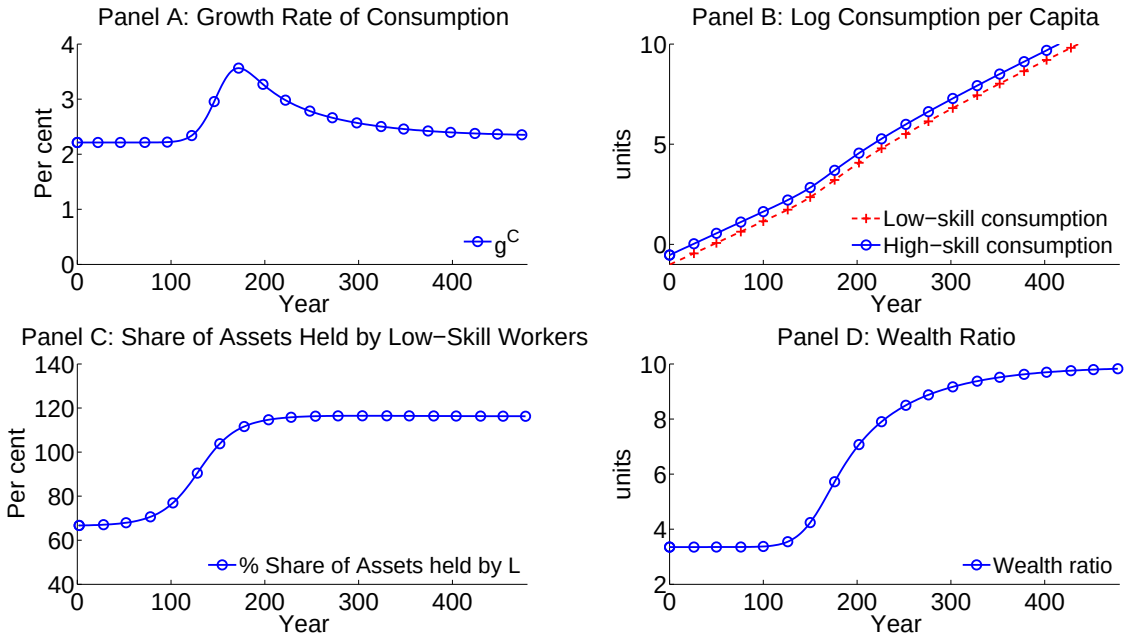

Figure 12: Consumption and wealth for baseline parameters. Panel A shows yearly growth rates for consumption, Panel B log consumption of high-skill workers and low-skill workers (per capita), Panel C the share of assets held by low-skill workers and Panel D the wealth to GDP ratio.

The accumulation of asset holdings by low-skill workers predicted by the model seems counter-factual, it results from our assumptions of infinitely lived agents with identical discount rates and no financial constraints. Reversing these unrealistic assumptions would change the evolution of the consumption side of the economy but should not alter the main results which are about the production side.

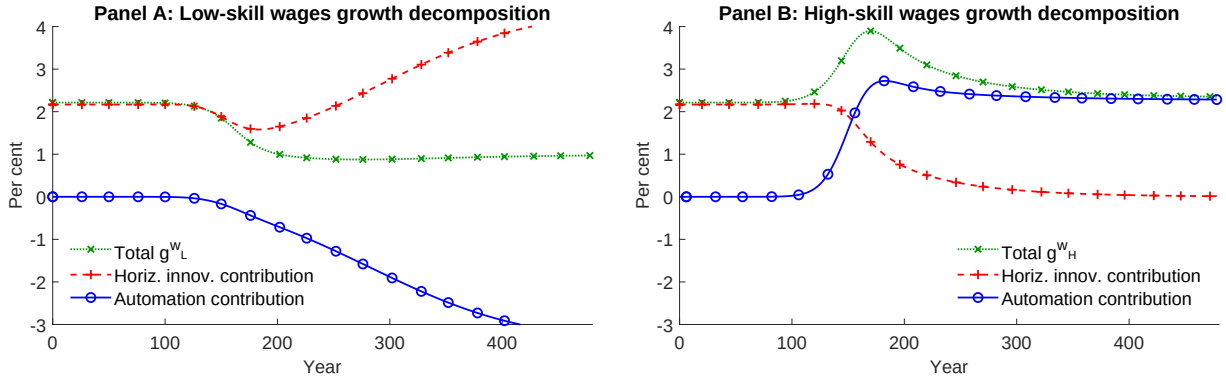

Figure 13: Growth decomposition. Panel A: The growth rate of low-skill wages and the instantaneous contribution from horizontal innovation and automation, respectively. Panel B is analogous for high-skill wages. See text for details.

Growth decomposition. Figure 13 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_L and w_H change?[43] 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 instantaneous 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. We stress that this growth decomposition captures the immediate effect of automation or horizontal innovation. This should not be interpreted as "automation being harmful" to low-skill workers in general. In fact, as we demonstrate in Section 3.4, an increase in the effectiveness of the automation technology, η, will have positive long-term consequences. A decomposition of g_t^{GDP} would look similar to the decomposition of $g_t^{w_H}$: while instantaneous growth is initially almost entirely driven by horizontal innovation, automation becomes increasingly important in explaining it (long-run growth, however, is ultimately determined by the endogenous rate of horizontal innovation).

7.6.2 Negative growth for low-skill wages

This section presents two examples with negative growth for low-skill wages. We ensure temporary negative growth in low-skill wages in Figure 14 by setting $\tilde{\kappa} = 0.49$, thereby

[43] More specifically we can write $w_{Lt} = f(N_t, G_t, H_t^P)$, using equations (7) and (6). Differentiating with respect to time and using equation (31) gives:

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

Figure 13 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 w_{Ht}.

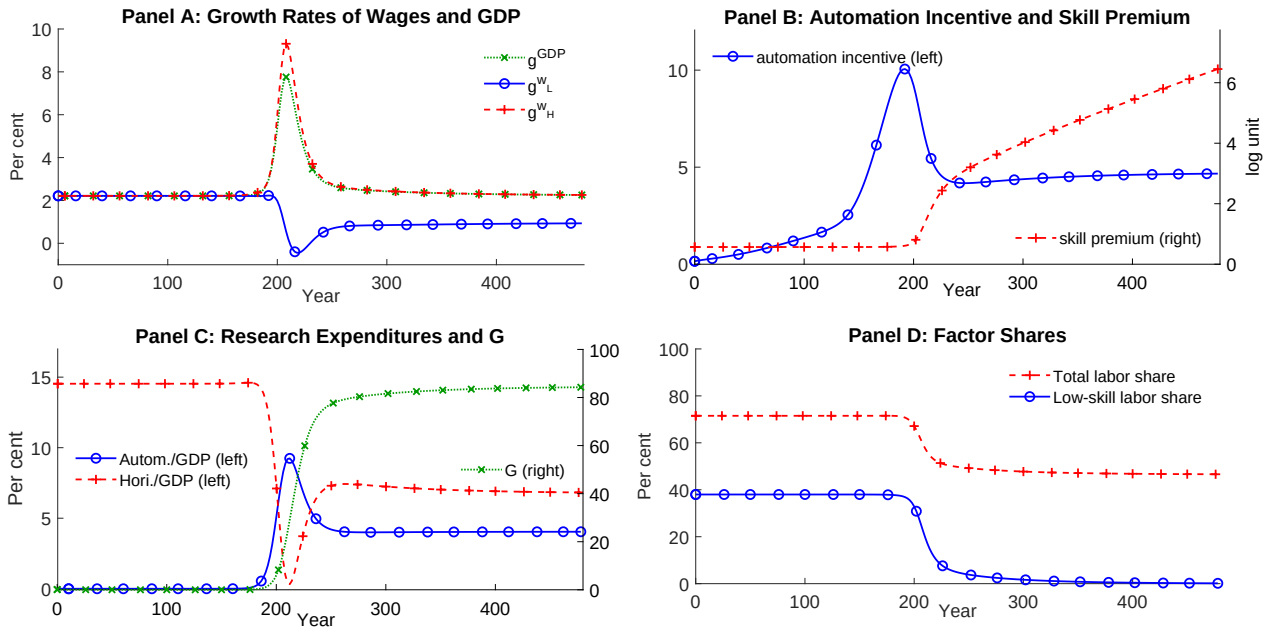

Figure 14: Transitional Dynamics with temporary decline in low-skill wages with an automation externality. Note: same as for Figure 11 but with an automation externality of $\tilde{\kappa} = 0.49$.

introducing the externality in automation. Initially, G_t is small and the automation technology is quite unproductive. Hence, Phase 2 starts later, even though the ratio $\left(V_t^A - V_t^N\right) / \left(w_{Ht}/N_t\right)$ 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_{Lt} (and increase in high-skill wage w_{Ht}) lowers the incentive to automate (Panel B), which in return reduces automation. Note that in a model with endogenous adaptation where automation involves the payment of a fixed cost every period instead of a R&D sunk cost, it would not be possible to obtain even a temporary decline in low-skill wages as firms would stop paying the fixed cost as soon as wages decline. Here, the discounted value of the difference in profits between automated and non-automated firms stays at a high level throughout ensuring that automation nonetheless remains at a high level.

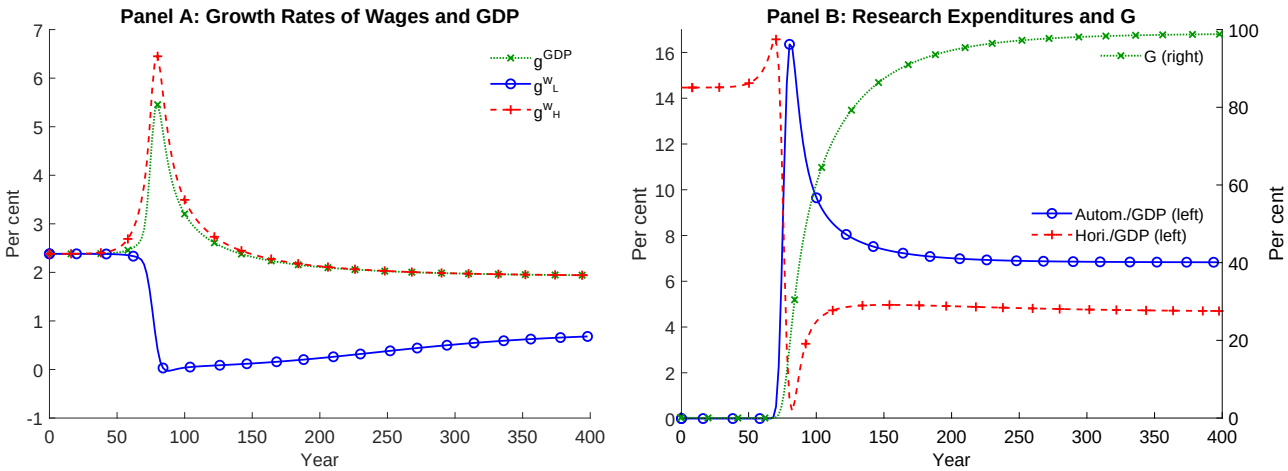

Figure 15: Transitional dynamics with temporary decline in low-skill wages without an automation externality.

Low-skill wages can also drop for $\tilde{\kappa} = 0$ as shown in Figure 15 where low-skill wages slightly decline for a short time period. The associated parameters are given in Table 5. The crucial parameter change is an increase in κ, such that the automation technology is less concave. This delays Phase 2, which is then more intense and leads to a sharp increase in high-skill wages, reducing considerably horizontal innovation.

Table 5: Baseline Parameter Specification

σ	ϵ	β	H	L	θ	η	κ	$\tilde{\varphi}$	ρ	$\widetilde{\kappa}$	γ
3	73	0.72	0.35	0.65	2	0.2	0.97	0.25	0.022	0	0.28

7.6.3 The effect of the innovation parameters

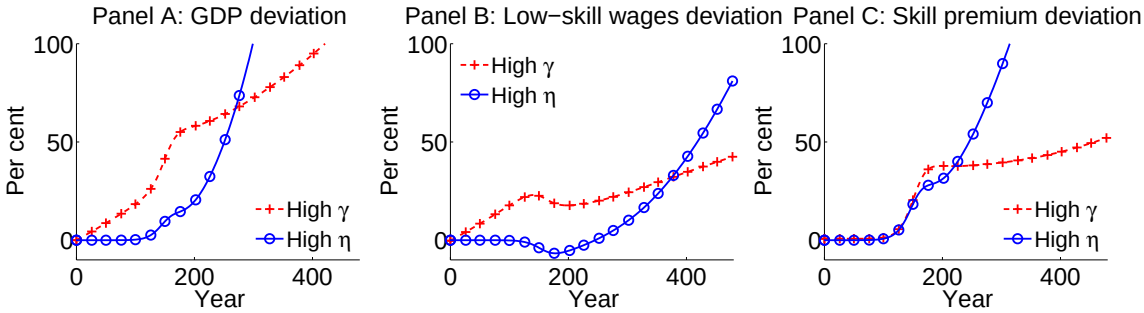

Figure 16: Deviations from baseline model for more productive horizontal innovation technology (γ) and more productive automation technology (η).

Figure 16 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. 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.

7.6.4 Systematic comparative statics

In this section we carry a systematic comparative exercise with respect to the parameters of the model, namely $\sigma, \epsilon, \beta, \rho, \theta, \widetilde{\varphi}, \eta, \kappa, \tilde{\kappa}, \gamma, H/L$ (we keep $H + L = 1$), N_0, G_0. We show the evolution of the growth rate of high-skill and low-skill wages and the share of automated products for the baseline parameters and two other values for one parameter, keeping all the other ones fixed. In all cases, the broad structure of the transitional dynamics in three phases is maintained.

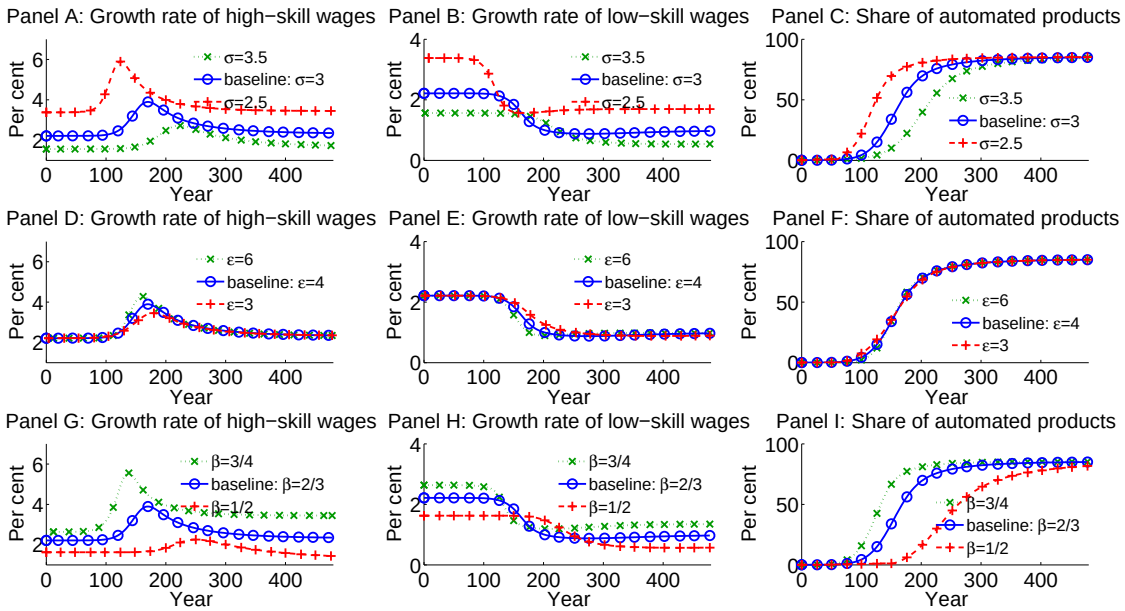

Figure 17: Comparative statics with respect to the elasticity of substitution across products (σ), the elasticity of substitution between machines and low-skill workers in automated firms (ϵ) and the factor share of low-skill workers and machines in production (β).

Figures 17.A,B,C show that a higher elasticity of substitution across products σ reduces the growth rate of the economy (the elasticity of output with respect to the number of products is lower), which leads to a delayed transition. The asymptotic growth rate of low-skill wages is a smaller fraction of that of high-skill wages (following Proposition 2), since automated products are a better substitute for non-automated ones. Figures 17.D,E,F show that the elasticity of substitution between machines and low-skill workers in automated firms, ϵ, plays a limited role (as long as $\mu < 1$), a higher elasticity reduces the growth of low-skill wages and increases that of high-skill wages during Phase 2. Figures 17.G,H,I show that a lower factor share in production for high-skill workers (a higher β) increases the growth rate of the economy (high-skill wages are lower which favors innovation). As a result, Phase 2 occurs sooner. Besides, following Proposition 2, the asymptotic growth rate of low-skill wages is a lower fraction of that of high-skill wages (the cost advantage of automated firms being larger).

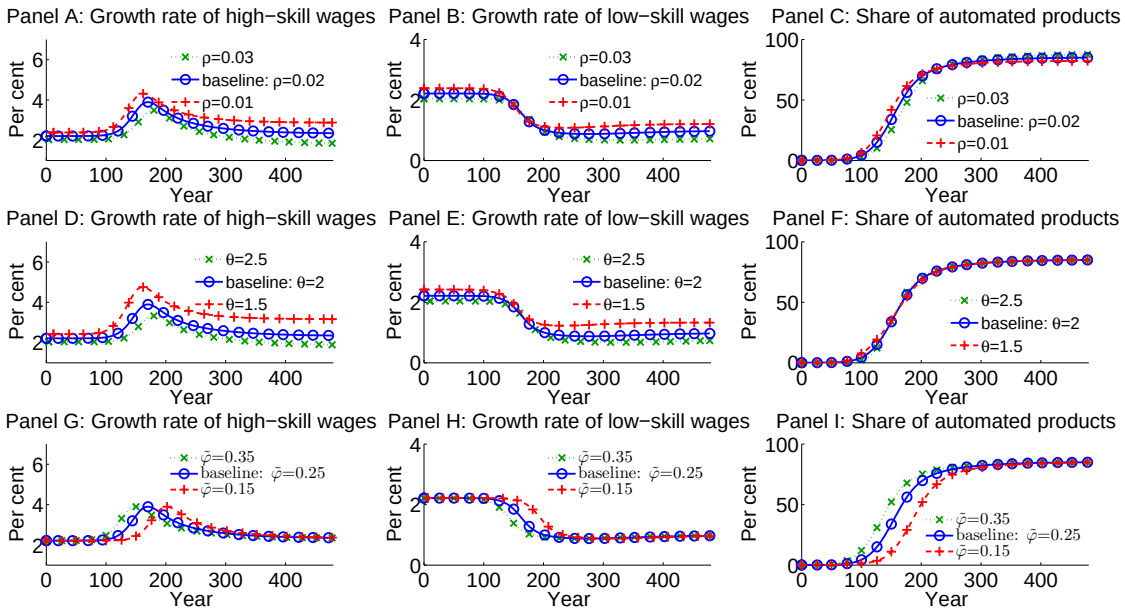

Figure 18: Comparative statics with respect to the discount rate (ρ), the inverse elasticity of intertemporal substitution (θ) and the productivity of machines ($\tilde{\varphi}$)

Figures 18.A,B,C show that a higher discount rate ρ reduces the growth rate of the economy, which slightly postpones Phase 2. At the time of Phase 2, the growth rate of low-skill wages is not affected much by the discount rate: on one hand, since low-skill wages are lower Phase 2 is postponed, which favor low-skill wages' growth, but on the other hand, horizontal innovation is lower which negatively affects low-skill wages. A lower elasticity of intertemporal substitution (a higher θ) has a similar effect on the economy's growth rate (Figures 18.D,E,F). Figures 18.G,H,I show that the productivity of machines ($\widetilde{\varphi}$) only affects the timing of Phase 2 (Phase 2 occurs sooner when machines are more productive).

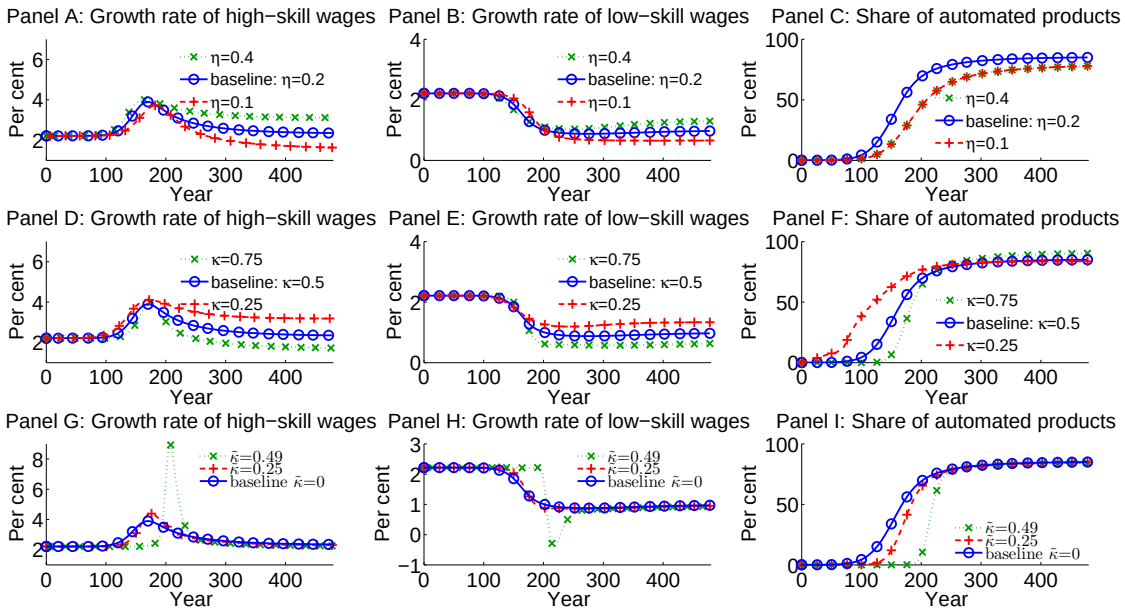

Figure 19: Comparative statics with respect to the automation productivity (η), the concavity of the automation technology (κ) and the automation externality ($\tilde{\kappa}$)

The comparative statics with respect to the automation technology shown in Figures 19.A,B,C follow the pattern described in the text. A less concave automation technology (higherκ) delays Phase 2 and reduces the economy's growth rate. It particularly affects the growth rate of low-skill wages in Phase 2 (as the increase in automation expenses comes more at the expense of horizontal innovation)—see Figures 19.D,E,F. The role of the automation externality has already been discussed in the text, Figures 19.G,H,I reveal that for a mid-level of the automation externality ($\tilde{\kappa} = 0.25$), the economy looks closer to the economy without the automation externality than to the economy with a large automation externality.

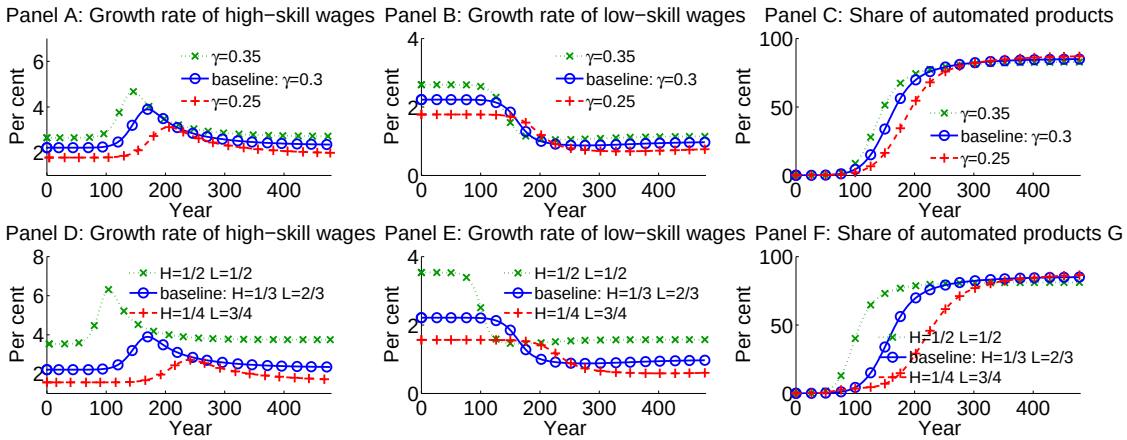

Figure 20: Comparative statics with respect to the horizontal innovation productivity (η) and the skill ratio (H/L)

Figures 20.A,B,C show the impact of the horizontal innovation parameter γ, which was already discussed in the text. Figures 20.D,E,F show that a higher ratio H/L naturally leads to a higher growth rate, which implies that Phase 2 occurs sooner. Figures 21.A,B,C show that a higher initial number of products simply advance the entire evolution of the economy. Figures 21.D,E,F show that a higher initial value for the share of automated products (even as high as the steady-state value G^*) barely affects the evolution of the economy, the share of automated products initially drops quickly as there is little automation to start with.

7.7 Simulation technique

In the following we describe the simulation techniques employed in Appendix 7.6 for the baseline model presented in 2. The approach for the extensions and the quantitative exercise of Section 4 follow straightforwardly. Let $\mathbf{x}_t \equiv (n_t, G_t, \hat{h}_t^A, \chi_t, \omega_t)$ and note that equation (47) defines ω implicitly. We can therefore write equations (30), (31), (37) and (38) as a system of autonomous differential equations $(\dot{n}_t, \dot{G}_t, \dot{\hat{h}}_t^A, \dot{\chi}_t) = F(\mathbf{x}_t)$ with initial conditions on state variables as (n_0, G_0) and an auxiliary equation of $\omega_t = \vartheta(\mathbf{x}_t)$.

For the numerical solution, we specify a (small) time period of $dt > 0$ and a (large) number of time periods T. Using this we approximate the four differential equations by $(T-1) \times 4$ errors as:

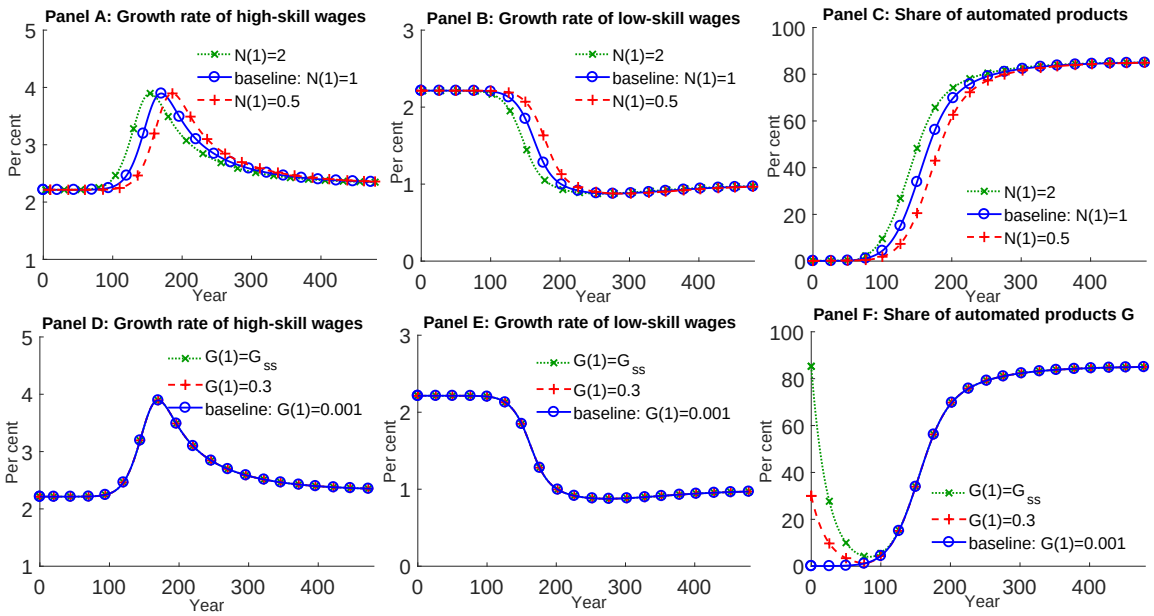

Figure 21: Comparative statics with respect to the initial number of products N_0 and the initial share of automated products G_0

$$s_t = (\frac{n_{t+1} - n_t}{dt}, \frac{G_{t+1} - G_t}{dt}, \frac{\hat{h}^A_{t+1} - \hat{h}^A_t}{dt}, \frac{\chi_{t+1} - \chi_t}{dt}) - F((\mathbf{x}_t + \mathbf{x}_{t+1})/2),\ t = \{1, ...T-1\}$$

with T corresponding errors for ω_t :

$$s^\omega_t = \omega_t - \vartheta(\mathbf{x}_t),\ t = \{1, ..., T\}.$$

Following Lemma 1, for a set of parameter values, the system admits an asymptotic steady state. We assume that the system has reached this asymptotic steady state by time T and restrict $\hat{h}^A_T$ and χ_T accordingly. Together with the initial conditions ($n_1 = n^{start}$ and $G_1 = G^{start}$) this leads to a vector of errors:

$$s_T \equiv (n_1 - n^{start}, G_1 - G^{start}, \hat{h}^A_T - \hat{h}^{A*}, \chi_T - \chi^*)'.$$

Letting $\mathbf{x} = \{\mathbf{x}_t\}^T_{t=1}$, we then stack errors to get a vector, $S(\mathbf{x})$, of length $5T$ and solve the following problem:

$$\hat{\mathbf{x}} = argmin_{\mathbf{x}} S(\mathbf{x})' W S(\mathbf{x}),$$

for a $5T \times 5T$ diagonal weighting matrix, W, and the $5T$ vector $\mathbf{x}$. For $dt \to 0$ and

$T \to \infty$ $S(\mathbf{x})'WS(\mathbf{x}) \to 0$. For the simulations we set $dt = 2$ and $T = 2000$. We accept the solution when $S(\hat{\mathbf{x}})'WS(\hat{x}) < 10^{(-7)}$, but the value is typically less than $10^{(-20)}$. The choice of weighting matrix matters somewhat for the speed of convergence, but is inconsequential for the final result. With the solution $\{\hat{\mathbf{x}}_t, \hat{\omega}_t\}_{t=1}^{T}$ in hand, it is straightforward to find remaining predicted values.

7.8 Social planner problem

This section presents the solution to the social planner problem. After having set-up the problem, we derive the optimal allocation, emphasizing in particular the different inefficiencies in our competitive equilibrium. Then, we show the optimal allocation for our baseline parameters. Finally, we derive how the optimal allocation can be decentralized.

7.8.1 Characterizing the optimal allocation

We introduce the following notations: N_t^A (respectively N_t^N) denotes the mass of automated (respectively non-automated) firms, L_t^A (respectively L_t^N) is the mass of low-skill workers hired in automated (respectively non-automated) firms, and $H_t^{P,A}$ (respectively $H_t^{P,N}$) is the mass of high-skill workers hired in production in automated (respectively non-automated) firms. The social planner problem can then be written as (we write the Lagrange multipliers next to each constraint):

$$\max \int_0^\infty e^{-\rho t} \frac{C_t^{1-\theta}}{1-\theta}$$

such that

$$\widetilde{\lambda}_t : C_t + X_t = F\left(L_t^A, H_t^{P,A}, X_t, L_t^N, H_t^{P,N}, N_t^A, N_t^N\right),$$

with

$$F \equiv \left(\begin{array}{c} \left(N_t^A\right)^{\frac{1}{\sigma}} \left(\left(\widetilde{\varphi} X_t^{\frac{\epsilon-1}{\epsilon}} + \left(L_t^A\right)^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}\beta} \left(H_t^{P,A}\right)^{1-\beta} \right)^{\frac{\sigma-1}{\sigma}} \\ + \left(N_t^N\right)^{\frac{1}{\sigma}} \left(\left(L_t^N\right)^{\beta} \left(H_t^{P,N}\right)^{1-\beta} \right)^{\frac{\sigma-1}{\sigma}} \end{array} \right)^{\frac{\sigma}{\sigma-1}},$$

$$\widetilde{w}_t : L_t^A + L_t^N = L,$$

$$\widetilde{v}_t : H_t^{P,A} + H_t^{P,N} + H_t^A + H_t^D = H,$$

$$\widetilde{\zeta}_t : \dot{N}_t^N = \gamma \left(N_t^A + N_t^N\right) H_t^D - \eta \left(N_t^A\right)^{\widetilde{\kappa}} \left(N_t^N + N_t^A\right)^{\kappa-\widetilde{\kappa}} \left(H_t^A\right)^{\kappa} \left(N_t^N\right)^{1-\kappa},$$

$$\widetilde{\xi}_t : \dot{N}_t^A = \eta \left(N_t^A\right)^{\widetilde{\kappa}} \left(N_t^N + N_t^A\right)^{\kappa-\widetilde{\kappa}} \left(H_t^A\right)^{\kappa} \left(N_t^N\right)^{1-\kappa},$$

$$H_t^D \geq 0.$$

The first order condition with respect to C_t gives

$$C_t^{-\theta} = \widetilde{\lambda}_t$$

To denote the ratio of the Lagrange parameter of each constraint with respect to $\widetilde{\lambda}_t$ (that is the shadow value expressed in units of final good at time t), we remove the tilde (hence $w_{Lt} \equiv \widetilde{w}_{Lt}/\widetilde{\lambda}_t$ is the shadow wage of low-skill workers).

The first order conditions with respect to X_t implies that

$$\frac{\partial F}{\partial X_t} = 1, \tag{121}$$

so that the shadow price of a machine must be equal to 1. First order conditions with respect to L_t^A, L_t^N, $H_t^{P,A}$, $H_t^{P,N}$ lead to

$$w_{Lt} = \frac{\partial F}{\partial L_t^A} = \frac{\partial F}{\partial L_t^N} \text{ and } w_{Ht} = \frac{\partial F}{\partial H_t^{P,A}} = \frac{\partial F}{\partial H_t^{P,N}}, \tag{122}$$

so that labor inputs are paid their marginal product in aggregate production. This is not the case in the competitive equilibrium, where labor inputs are paid their marginal product but products are priced with a mark-up as they are provided by a monopolist. It is easy to show that for a given H_t^P, the optimal provision of machines and allocation of high-skill and low-skill workers across firms can be obtained if the purchase of all products is subsidized by at rate $1/\sigma$ (a lump-sum tax finances the subsidy).

The first-order conditions with respect to N_t^N and N_t^A are given by:

$$\rho\widetilde{\zeta}_t - \dot{\widetilde{\zeta}}_t = \begin{array}{c} \widetilde{\lambda}_t \frac{\partial F}{\partial N_t^N} + \widetilde{\zeta}_t \gamma H_t^D + \left(\widetilde{\xi}_t - \widetilde{\zeta}_t\right) \eta \left(H_t^A\right)^{\kappa} \left(N_t^N\right)^{-\kappa} \\ \times \left(N_t^A\right)^{\widetilde{\kappa}} \left(\left(1-\widetilde{\kappa}\right) N_t^N + \left(1-\kappa\right) N_t^A\right) \left(N_t^N + N_t^A\right)^{\kappa-\widetilde{\kappa}-1} \end{array}, \tag{123}$$

$$\rho\widetilde{\xi}_t - \dot{\widetilde{\xi}}_t = \begin{array}{c} \widetilde{\lambda}_t \frac{\partial F}{\partial N_t^A} + \widetilde{\zeta}_t \gamma H_t^D + \left(\widetilde{\xi}_t - \widetilde{\zeta}_t\right) \eta \left(H_t^A\right)^{\kappa} \\ \times \left(N_t^N\right)^{1-\kappa} \left(N_t^A\right)^{\widetilde{\kappa}-1} \left(\widetilde{\kappa} N_t^N + \kappa N_t^A\right) \left(N_t^N + N_t^A\right)^{\kappa-\widetilde{\kappa}-1} \end{array}. \tag{124}$$

Interestingly, $\frac{\partial F_t}{\partial N_t^N}$ and $\frac{\partial F}{\partial N_t^A}$ correspond to the profits realized by a non-automated and an automated firm respectively in the equilibrium once the subsidy to the use of products

is implemented. Therefore we denote

$$\pi_t^N = \frac{\partial F_t}{\partial N_t^N} \text{ and } \pi_t^A = \frac{\partial F_t}{\partial N_t^A}$$

Further the (shadow) interest rate is given by $r_t = \rho + \theta \frac{\dot{C}_t}{C_t} = \rho - \frac{\dot{\lambda}_t}{\lambda_t}$. Using that $H_t^A = (1 - G_t) N_t h_t^A$, we can rewrite (123) and (124) as:

$$r_t \zeta_t = \pi_t^N + \zeta_t g_t^N + (\xi_t - \zeta_t) \eta (G_t)^{\widetilde{\kappa}} N_t^{\kappa} \left(h_t^A\right)^{\kappa} ((1 - \widetilde{\kappa})(1 - G_t) + (1 - \kappa) G_t) + \dot{\zeta}_t, \quad (125)$$

$$r_t \xi_t = \pi_t^A + \zeta_t g_t^N + (\xi_t - \zeta_t) \eta (G_t)^{\widetilde{\kappa}} N_t^{\kappa} \left(h_t^A\right)^{\kappa} (1 - G_t) \left(\widetilde{\kappa} \frac{1 - G_t}{G_t} + \kappa\right) + \dot{\xi}_t. \quad (126)$$

These expressions parallel equations (14) and (15) in the paper. The rental social value of a non-automated firm ($r_t \zeta_t$) consists of the current value of one product (which equals the profits when the optimal subsidy to the use of intermediate products is in place), its positive impact on the horizontal innovation technology (the productivity of which is γN_t), its positive impact on the automation technology (which results from the direct externality embedded in the automation technology from the number of firms diminished by the additional externality coming from the share of automated products), the expected increase in its value if it becomes automated minus the cost of the resources required (the difference between these two terms is positive since the automation technology is concave) and the change in its value. The rental social value of an automated firms ($r_t \xi_t$) is the sum of the profits, its impact on horizontal innovation (through the same externality as non-automated firm), its impact on the automation technology (which results from two externalities as both the number of firms and the share of automated products improve the automation technology), and the change in its value.

The first order condition with respect to H_t^D gives (together with $H_t^D \geq 0$):

$$w_{Ht} \geq \zeta_t \gamma N_t, \quad (127)$$

with equality when $H_t^D > 0$. This equation is the counterpart of (17) in the equilibrium case, it stipulates that when horizontal innovation takes place the social value of a non-automated product equals the cost of creating one. The first-order condition with respect to H_t^A gives:

$$w_{Ht} = (\xi_t - \zeta_t) \kappa \eta (G_t)^{\widetilde{\kappa}} N_t^{\kappa} \left(h_t^A\right)^{\kappa - 1}. \quad (128)$$

This equation is the counterpart of (16) in the equilibrium case. Everything else given, $\xi_t - \zeta_t$ increases with $\pi_t^A - \pi_t^N$, which increases with w_{Lt}, therefore this equation shows that automation increases with low-skill wages (everything else given), just as in the equilibrium case.

7.8.2 System of differential equations and steady state

After having introduced the same variables as in the equilibrium case, one can follow the same steps and derive a system of differential equation in $\left(n_t, G_t, \widehat{h}_t^A, \chi_t\right)$ which characterizes the solution (when there is positive growth). Equations (30) and (31) still hold, while equations (37) and (38) are replaced with

$$\dot{\widehat{h}}_t^A = \frac{\frac{\gamma \widehat{h}_t^A}{1-\kappa}\left(\omega_t n_t \left(\varphi + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu} \frac{\widehat{\pi}_t^A}{\widehat{v}_t} + \frac{1-\kappa+(\kappa-\widetilde{\kappa})(1-G_t)}{\kappa}\widehat{h}_t^A\right)}{-\frac{\eta\kappa G_t^{\widetilde{\kappa}}}{1-\kappa}\left(\widehat{h}_t^A\right)^{\kappa}\left(1-\omega_t n_t\left(\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu}\right)\frac{\widehat{\pi}_t^A}{\widehat{v}_t} + \eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa+1} + \frac{1-\widetilde{\kappa}}{1-\kappa} g_t^N \widehat{h}_t^A}, \tag{129}$$

$$\dot{\chi}_t = \chi_t\left(\gamma\omega_t n_t\left(\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu}\frac{\widehat{\pi}_t^A}{\widehat{v}_t} + \gamma\frac{1-\kappa+(\kappa-\widetilde{\kappa})(1-G_t)}{\kappa}\widehat{h}_t^A - \rho - (\theta-1)\psi g_t^N\right).$$

g_t^N is still given by (46), $\frac{\widehat{\pi}_t^A}{\widehat{v}_t}$, H_t^P and ω_t are now given by

$$\frac{\widehat{\pi}_t^A}{\widehat{v}_t} = \frac{\psi\left(\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} H_t^P}{G_t\left(\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + (1-G_t)\,\omega_t n_t},$$

$$H_t^P = \frac{(1-\beta)^{\frac{1}{\theta}}\,\beta^{\frac{\beta}{1-\beta}\left(\frac{1}{\theta}-1\right)}\chi_t^{\frac{1}{\theta}}\left(G_t\left(\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + (1-G_t)\,\omega_t n_t\right)^{\psi\left(\frac{1}{\theta}-1\right)+1}}{G_t\left((1-\beta)\,\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)\left(\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu-1} + (1-G_t)\,\omega_t n_t},$$

$$\omega_t = \left(\begin{array}{c}\beta^{\frac{1}{1-\beta}}\frac{H_t^P}{L}\left(G_t\left(\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu-1}(\omega_t n_t)^{\frac{1-\mu}{\mu}} + (1-G_t)\right) \\ \times\left(G_t\left(\varphi+(\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + (1-G_t)\,\omega_t n_t\right)^{\psi-1}\end{array}\right)^{\frac{\beta(1-\sigma)}{1+\beta(\sigma-1)}},$$

which replace (43), (45) and (47).

One can then solve for a steady state of this system with $G^* > 0$ (and $\left(g^N\right)^* > 0$ so

that $n^* = 0$). (49) and (51) still apply, but (50) is replaced with

$$\widehat{h}^{A*} = \frac{\kappa}{\gamma} \frac{\rho + (\theta - 1)\,\psi g^{N*}}{1 - \kappa + (1 - G^*)\,(\kappa - \widetilde{\kappa})}, \tag{130}$$

and (53) with

$$f^{sp}\left(g^{N*}\right) \equiv \frac{\rho + (\theta - 1)\,\psi g^{N*}}{\psi H^{P*}} \left(\frac{\left(\widehat{h}^{A*}\right)^{1-\kappa}}{\eta\kappa\,(G^*)^{\widetilde{\kappa}-1}} + \frac{1}{\gamma} \right),$$

which is obtained by fixing $\dot{\widehat{h}}_t^A = 0$ in (129) using (49) and (130). For g^{N*} large enough (but finite—and, in particular smaller than γH), H^{P*} is arbitrarily small, while for the same value G^* and $\widehat{h}^{A*}$ are bounded below and above. As before, this establishes that for g^{N*} large enough $f^{sp}\left(g^{N*}\right) > 1$. Furthermore $f^{sp}(0) = f(0)$, therefore condition 48 is also a sufficient condition for the existence of a steady state with positive growth and $G^* > 0$ for the system of differential equations.

7.8.3 Decentralizing the optimal allocation

We have already seen that the "static" optimal allocation given H_t^P is identical to the equilibrium allocation once a subsidy to the use of products $1/\sigma$ is in place. The "dynamic" part of the problem consists of the allocation of high-skill workers across the two types of innovation and production. Therefore, we postulate that a social planner can decentralize the optimal allocation using the subsidy to the use of intermediate products and subsidies (or taxes) for high-skill workers hired in automation $\left(s_t^A\right)$ and in horizontal innovation $\left(s_t^H\right)$. Let us consider such an equilibrium and introduce the notations $\Omega_t^A \equiv 1 - s_t^A$ and Ω_t^H similarly defined. In this situation, the law of motion for the private value of an automated firm, V_t^A, is still given by (14), for a non-automated firm it obeys:

$$r_t V_t^N = \pi_t^N - \Omega_t^A w_{Ht} h_t + \eta\,(G_t)^{\widetilde{\kappa}}\, N_t^{\kappa} \left(h_t^A\right)^{\kappa} \left(V_t^A - V_t^N\right) + \dot{V}_t^N, \tag{131}$$

instead of (15), the first-order condition for automation is given by:

$$\kappa\eta\,(G_t)^{\widetilde{\kappa}}\, N_t^{\kappa} \left(h_t^A\right)^{\kappa-1} \left(V_t^A - V_t^L\right) = \Omega_t^A w_{Ht}, \tag{132}$$

instead of (16), while the free entry condition, when $g_t^N > 0$, is given by

$$\gamma N_t V_t^N = \Omega_t^H w_{Ht}, \tag{133}$$

instead of (17). For Ω_t^A and Ω_t^H to decentralize the optimal allocation it must be that these 4 equations hold together with (125), (126), (127) and (128).

Using (127) and (133), we then get that Ω_t^H must satisfy

$$\Omega_t^H \zeta_t = V_t^N, \tag{134}$$

similarly, using (128) and (132), we get

$$\Omega_t^A (\xi_t - \zeta_t) = V_t^A - V_t^L. \tag{135}$$

Plugging (134) and (135) in (131), we get that

$$r_t \zeta_t = \frac{\pi_t^N}{\Omega_t^H} - \frac{\Omega_t^A}{\Omega_t^H} w_{Ht} h_t + \eta (G_t)^{\widetilde{\kappa}} N_t^{\kappa} \left(h_t^A\right)^{\kappa} \frac{\Omega_t^A}{\Omega_t^H} (\xi_t - \zeta_t) + \frac{\dot{\Omega}_t^H}{\Omega_t^H} \zeta_t + \dot{\zeta}_t. \tag{136}$$

Similarly, using (135) and the difference between (14) and (131) gives:

$$r_t (\xi_t - \zeta_t) = \frac{\pi_t^A - \pi_t^N}{\Omega_t^A} + w_{Ht} h_t - \eta (G_t)^{\widetilde{\kappa}} N_t^{\kappa} \left(h_t^A\right)^{\kappa} (\xi_t - \zeta_t) + \frac{\dot{\Omega}_t^A}{\Omega_t^A} (\xi_t - \zeta_t) + \dot{\xi}_t - \dot{\zeta}_t. \tag{137}$$

Combining (136) with (125), using (128) and (127) and the definition of Ω_t^A and Ω_t^H, we get:

$$\dot{s}_t^H = \begin{array}{c} \frac{\gamma \widehat{\pi}_t^N}{\widehat{v}_t} s_t^H - \left(1 - s_t^H\right) g_t^N \\ + \frac{\gamma \widehat{h}_t^A}{\kappa} \left(\left(1 - s_t^A\right)(1 - \kappa) + \left(1 - s_t^H\right)\left(\widetilde{\kappa}(1 - G_t) + \kappa G_t - 1\right)\right) \end{array}. \tag{138}$$

Similarly combining (137) with the difference between (126) and (125) and using (127) gives:

$$\dot{s}_t^A \frac{\left(\widehat{h}_t^A\right)^{1-\kappa}}{\eta G_t^{\widetilde{\kappa}}} = \kappa \frac{\widehat{\pi}_t^A - \widehat{\pi}_t^N}{\widehat{v}_t} s_t^A - \widetilde{\kappa} \left(1 - s_t^A\right) \widehat{h}_t^A \frac{1 - G_t}{G_t}. \tag{139}$$

Therefore, in steady state, we have

$$s_{\infty}^{A} = \frac{\widetilde{\kappa}\widehat{h}_{\infty}^{A}\left(1-G_{\infty}\right)}{\kappa\psi H_{\infty}^{P} + \widetilde{\kappa}\widehat{h}_{\infty}^{A}\left(1-G_{\infty}\right)} \geq 0.$$

Note from (139) that the share of automated products, s_t^A, must always be non-negative, otherwise it cannot converge to a positive value, therefore $s_t^A \geq 0$ everywhere (and in fact > 0 if $\widetilde{\kappa} \neq 0$). Furthermore, if $\widetilde{\kappa} = 0$, $s_t^A = 0$ everywhere, the only externality in automation comes from the total number of products, therefore the equilibrium features the optimal amount of automation investment (when the monopoly distortion is corrected and the optimal subsidy to horizontal innovation is implemented).

(138) gives the steady state value of the subsidy to horizontal innovation as:

$$s_{\infty}^{H} = 1 - \frac{\gamma\widehat{h}_{\infty}^{A}\left(1-\kappa\right)\left(1-s_{\infty}^{A}\right)}{\kappa g_{\infty}^{N} + \gamma\widehat{h}_{\infty}^{A}\left(1-\widetilde{\kappa}\left(1-G_{\infty}\right)-\kappa G_{\infty}\right)}.$$

In addition, knowing that $s_t^A \geqslant 0$, imposes that $s_t^H > 0$—as $s_t^H < 0$ would lead to $\dot{s}_t^H < 0$.

7.8.4 Transitional dynamics for the social planner case

Figure 22 plots the transitional dynamics for the optimal allocation in our baseline case (with $\widetilde{\kappa} = 0$) and in the case where $\widetilde{\kappa} > 0$ analyzed in Figure 14. As shown in Panel A and C, the economy also goes through three phases as a higher (shadow) low-skill wage leads to more automation over time and a transition from a small share to a high share of automated products. Relative to Figure 11.A and Figure 14.A, the overall dynamics look quite similar but the growth rates are higher in the social planner case, and the transition to phase 2 now happens roughly at the same time with and without the automation externality, while in the equilibrium it is considerably delayed in the presence of the externality (as, effectively, the productivity of the automation technology is initially very low). In both cases, the social planner maintains a positive subsidy to horizontal innovation. When $\widetilde{\kappa} = 0$ (without the automation externality), the subsidy to automation is 0, while when $\widetilde{\kappa} > 0$ there is a positive subsidy to automation, which is the largest in Phase 1. This subsidy explains why Phase 2 now starts at around the same time.

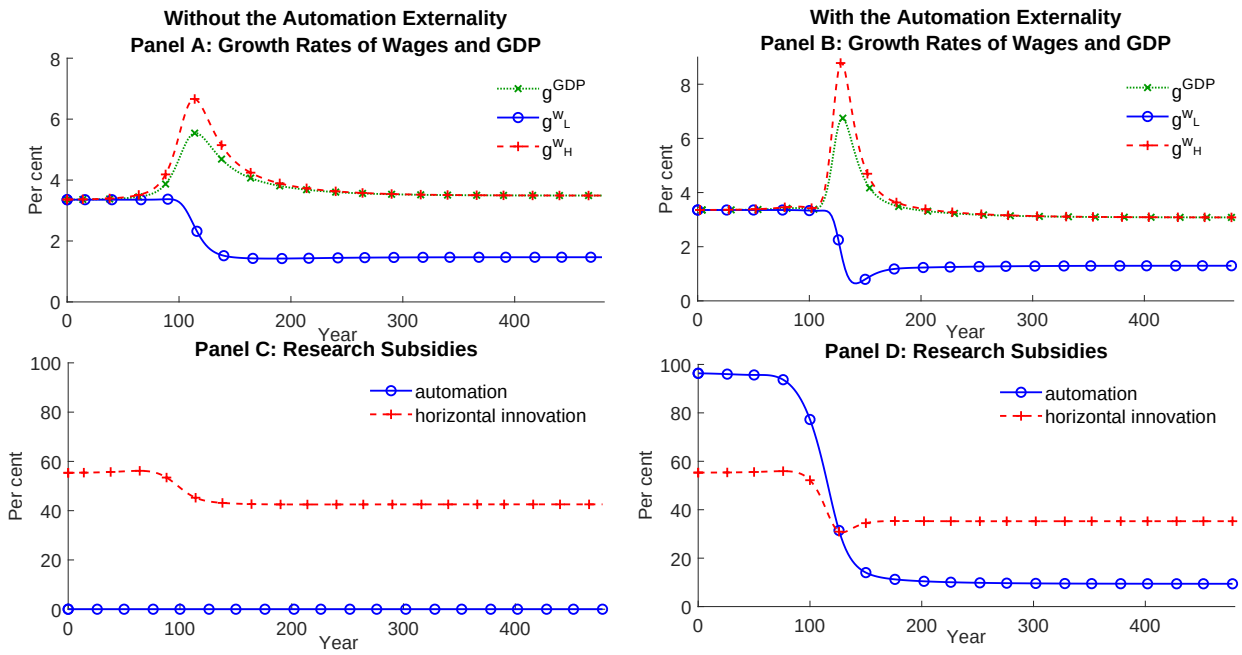

Figure 22: Transitional Dynamics in the Social Planner Case. Panel A and B, baseline case. Panel C and D, with $\tilde{\kappa} = 0.5$

7.9 An endogenous supply response in the skill distribution: dynamic model

In this Appendix we revisit the model with endogenous skill supply of Section 7.3 and we now characterize its behavior with endogenous innovation. We can derive the dynamic system as in the baseline model (see details below in Appendix 7.9.1). Lemma 1 can then be extended 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 $\bar{H}$.

Figure 23 shows the transitional dynamics for this model when the common parameters are the same as in Table 4, $\overline{H} = 1/3$ (so that $G^*, \hat{h}^{A*}, g^{N*}, \chi^*$ are the same as in the baseline model), $l = 1$ and $q = 0.3$. The figure looks similar to Figure 11, but the gap in steady-state between the low-skill growth rate and the high-skill growth rate is a bit smaller. In addition Panel B shows that the skill ratio increases from Phase 2 and Panel A shows that the growth rate is lower in Phase 1 as the mass of high-skill workers is lower then.

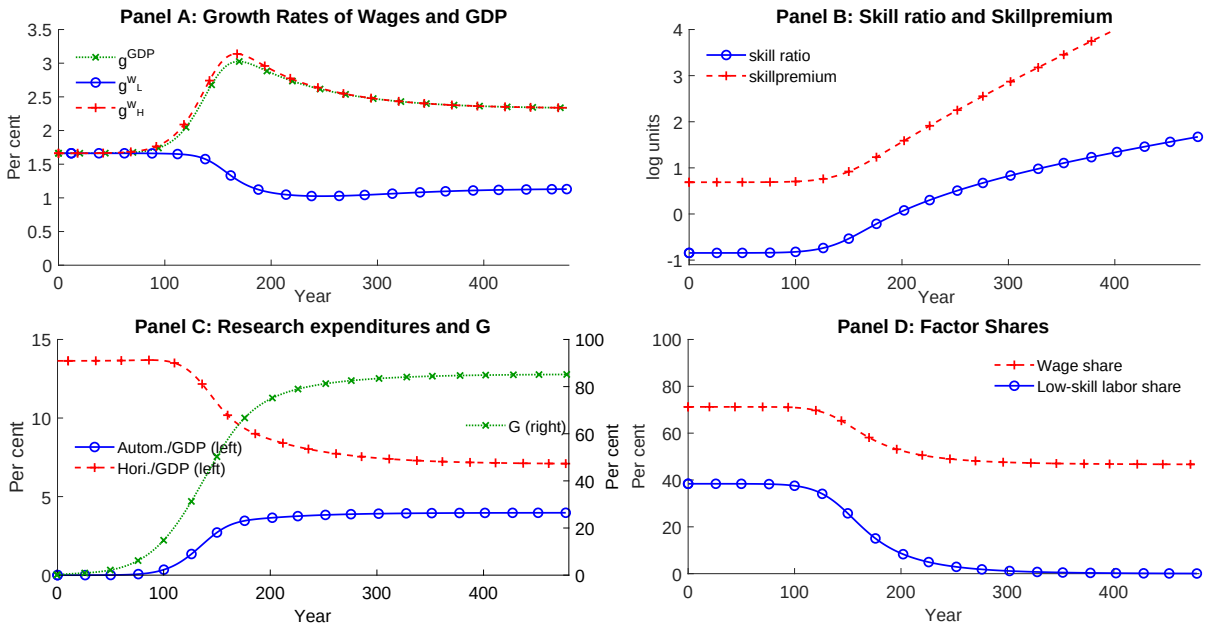

Figure 23: Transitional Dynamics for model with endogenous skill supply. Panel A shows growth rates for GDP, low-skill wages (w_L) and high-skill wages (w_H), Panel B the skill ratio and the skill premium, Panel C the total spending on horizontal innovation and automation as well as the share of automated products (G), and Panel D the wage share of GDP for total wages and low-skill wages.

7.9.1 Details on the dynamic system

It is convenient to redefine $n_t \equiv N_t^{\frac{-\beta}{(1-\beta)}\frac{(1+q)}{1+q+\beta(\sigma-1)}}$, we can then write the entire dynamic system as a system of differential equations in $\left(n_t, G_t, \widehat{h}_t^A, \chi_t\right)$ with two auxiliary variables ω_t and $\hat{j}_t \equiv \overline{j}_t n_t^{-\frac{q}{1+q}}$. Equations (30) is now given by

$$\dot{n}_t = -\frac{\beta}{1-\beta}\frac{1+q}{1+q+\beta(\sigma-1)} g_t^N n_t,$$

(31), (37), (38), (43), (45) still apply and equation (46) as well provided that H is replaced by H_t given by (75). ω_t is implicitly defined by:

$$\omega_t = \left(\begin{array}{c} \left(\frac{\sigma-1}{\sigma}\right)^{\frac{1+q}{1-\beta}} \beta^{\frac{1+\beta q}{1-\beta}} \left((1-\beta)\frac{1+q}{q}\right)^q \frac{H_t^P}{l^{1+q}\overline{H}} \left(G_t\left(1+\varphi\left(\omega_t n_t\right)^{-\frac{1}{\mu}}\right)^{\mu-1} + (1-G_t)\right) \\ \times \left(G_t\left(\varphi + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + (1-G_t)\,\omega_t n_t\right)^{\psi(1+q)-1} \end{array} \right)^{\frac{\beta(1-\sigma)}{1+q+\beta(\sigma-1)}},$$

which replaces (47) and is a rewriting of (79) and $\hat{j}_t$ is given by

$$\hat{j}_t = \left(\omega_t \frac{q}{1+q} \frac{\beta}{1-\beta} \frac{G_t \left(1 + \varphi \left(\omega_t n_t\right)^{-\frac{1}{\mu}}\right)^{\mu-1} + 1 - G_t}{G_t \left(\varphi + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + \left(1 - G_t\right) \omega_t n_t} \frac{H_t^P}{\overline{H}} \right)^{\frac{q}{1+q}}$$

which is derived using (76) and (78).

The steady state for this system involves $n^* = 0$ and therefore $\omega^*$and $\widehat{j}^*$ are positive constant (so that $\overline{j}^* = 0$: in steady-state all workers are high-skill). As a result $H^* = \overline{H}$, so that the steady state values of $\left(g^{N*}, G^*, \widehat{h}^{A*}, \chi^*\right)$ are identical to the baseline case with $\overline{H}$ replacing H.

7.10 Description of the Quantitative Model and Analytical Results

7.10.1 Set-up

To avoid repetitions, we already include the taxes of Section 7.11.6, namely, we assume that there is a tax τ_m on the rental rate of equipment and a tax τ_a on high-skill workers in automation innovations. The solution follows similar steps to the baseline case. We denote by $\widetilde{r}_t$ the gross rental rate of machines and by Δ their depreciation rate, such that:

$$\widetilde{r}_t = r_t + \Delta. \tag{140}$$

The Euler equation (18) still applies and the capital accumulation equation is given by (23). The unit cost of product i is now given by

$$c\left(w_L, w_H, \widetilde{r}, \alpha\left(i\right)\right) = \frac{\left(w_L^{1-\epsilon} + \alpha(i)\varphi\left(\widetilde{r}^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}\right)^{\frac{\beta_1}{1-\epsilon}} w_H^{\beta_2} \widetilde{r}^{\beta_3}}{\beta_1^{\beta_1} \beta_2^{\beta_2} \beta_3^{\beta_3}} \tag{141}$$

instead of (3) where $\varphi \equiv \widetilde{\varphi}^{\epsilon} \left(\beta_4^{\beta_4} \left(1 - \beta_4\right)^{1-\beta_4}\right)^{\epsilon-1}$. Define $\mu \equiv \beta_1 \left(\sigma - 1\right) / \left(\epsilon - 1\right)$, we can then derive the isocost curve as:

$$N^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma - 1} \frac{w_H^{\beta_2} \widetilde{r}^{\beta_3}}{\beta_1^{\beta_1} \beta_2^{\beta_2} \beta_3^{\beta_3}} \left(G \left(\varphi \left(\left(\left(1 + \tau_m\right) \widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right)^{\mu} + \left(1 - G\right) w_L^{\beta_1 (1-\sigma)} \right)^{\frac{1}{1-\sigma}} = 1. \tag{142}$$

The same steps as before allows us to obtain the relative demand for high-skill versus low-skill workers as:

$$\frac{w_H H^P}{w_L L} = \frac{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}\right)\left(\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} + \beta_2(1-G)\, w_L^{\beta_1(1-\sigma)}}{\beta_1\left(G w_L^{1-\epsilon}\left(\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu-1} + (1-G)\, w_L^{\beta_1(1-\sigma)}\right)}. \tag{143}$$

Similarly, taking the ratio of income going to high-skill workers in production over income going to machines owners, we obtain a relationship linking the gross rental rate of capital and high-skill wages:

$$\frac{\widetilde{r}K}{w_H H^P} = \frac{\left(\begin{array}{c} G\left(\beta_3 + \frac{\beta_1(1-\beta_4)\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{(1+\tau_m)\left(w_L^{1-\epsilon}+\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}\right)}\right)\left(\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} \\ +\beta_3(1-G)\, w_L^{\beta_1(1-\sigma)} \end{array}\right)}{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}\right)\left(\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} + \beta_2(1-G)\, w_L^{\beta_1(1-\sigma)}}. \tag{144}$$

7.10.2 Effect of technology on wages

First note that one can rewrite (143)

$$\frac{w_H H^P}{w_L L} = \frac{G\left(\beta_2 + \beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu} + \beta_2(1-G)}{\beta_1\left(G(\Phi+1)^{\mu-1} + (1-G)\right)}, \tag{145}$$

where we defined

$$\Phi \equiv \varphi\left(\frac{w_L}{((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}}\right)^{\epsilon-1} = \varphi\left(\left(\frac{w_L}{(1+\tau_m)\widetilde{r}}\right)^{1-\beta_4}\left(\frac{w_L}{w_H}\right)^{\beta_4}\right)^{\epsilon-1}. \tag{146}$$

In (145), the RHS is increasing in w_L and decreasing in w_H/w_L for given G, $\widetilde{r}$. Therefore, this equation defines the relative demand curve in the w_L, w_H space as rotating counter-clockwise (when $G > 0$) when w_L increases. Plugging (145) in (142) then defines w_L uniquely as a function of N, G, $\widetilde{r}$ and H^P. We can then derive the effect of changes in

G and N for given H^P and $\widetilde{r}$ (i.e. when K is perfectly elastically supplied) on wages, the skill premium and the labor share as follows:

Proposition 8. *Consider the equilibrium (w_L, w_H) determined by equations (145) and (142). Assume that $\epsilon < \infty$, 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 (w_H) and low-skill wages (w_L). Provided that $G > 0$, an increase in N also increases the skill premium w_H/w_L and decreases the labor share for $H \approx H^P$.

B) An increase in the share of automated products G (keeping N and H^P constant) increases the high-skill wages w_H, the skill premium w_H/w_L and decreases the labor share for $H \approx H^P$. Its impact on low-skill wages is ambiguous.

Proof. One can rewrite (142) as:

$$N^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma-1} \frac{\widetilde{r}^{\beta_3}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}} \left(\frac{w_H}{w_L}\right)^{\beta_2} w_L^{\beta_2} \left(\begin{array}{c} G\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} \\ +\left(1-G\right) w_L^{\beta_1(1-\sigma)} \end{array} \right)^{\frac{1}{1-\sigma}} = 1. \tag{147}$$

Using that (145) establishes $\frac{w_H}{w_L}$ as an increasing function of w_L otherwise independent of N, we get that (147) implies that w_L and therefore w_H/w_L (when $G > 0$) and w_H itself must increase in N.

(145) also establishes that $\frac{w_H}{w_L}$ increases in G for a given w_L. Therefore if w_L is increasing in G, then it is direct that $\frac{w_H}{w_L}$ and w_H both also increase in G. Assume on the contrary that w_L decreases in G, then in (142) the direct effect of an increase in G is to decrease the LHS (because $\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} > w_L^{\beta_1(1-\sigma)}$), in addition an increase in G would reduce w_L which further reduces the LHS. To maintain the inequality, it must be that w_H increases. Therefore in this case too, w_H increases in G and so does w_H/w_L.

This model is isomorphic to the previous one when $\beta_4 = \beta_3 = 0$ (with $\varphi\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\epsilon}$ replacing φ), therefore the impact of a change of G on w_L is also ambiguous.

The labor share is now given by

$$LS = \frac{w_L L + w_H H}{Y + \left(1+\tau_a\right) w_H \left(H - H^P\right)}.$$

As before profits are a share $\frac{1}{\sigma}$ of output so that

$$Y = \frac{\sigma}{\sigma - 1}\left(w_L L + w_H H^P + \widetilde{r}K + T_m\right), \tag{148}$$

where T_m denotes the tax proceeds from the tax on equipment. We have

$$\begin{aligned} &\frac{T_m}{w_H H^P} \\ &= \frac{G\frac{\tau_m \beta_1 (1-\beta_4)\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{(1+\tau_m)\left(w_L^{1-\epsilon}+\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}\right)}\left(\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu}}{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}\right)\left(\varphi\left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} + \beta_2 (1-G) w_L^{\beta_1(1-\sigma)}}. \end{aligned} \tag{149}$$

Then, we obtain:

$$LS = \frac{w_L L + w_H H}{\frac{\sigma}{\sigma-1}\left(w_L L + w_H H^P + \widetilde{r}K + T_m\right) + (1+\tau_a) w_H (H - H^P)}. \tag{150}$$

Assume that $H = H^P$, then we get that

$$LS = \frac{\sigma - 1}{\sigma}\left(1 + \frac{\widetilde{r}K + T_m}{w_L L + w_H H}\right)^{-1}.$$

Using (143), (144), (149) and (146), we obtain:

$$\begin{aligned} &\frac{\widetilde{r}K + T_m}{w_L L + w_H H} \\ &= \frac{G\left(\beta_3 + \beta_1 (1-\beta_4)\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu} + \beta_3 (1-G)}{G\left(\beta_2 + \beta_1\beta_4 + \beta_1\frac{(1-\beta_4)}{\Phi+1}\right)(\Phi+1)^{\mu} + (\beta_1+\beta_2)(1-G)} \\ &= \frac{\beta_3}{\beta_2+\beta_1} + \frac{1}{\beta_2+\beta_1}\frac{G(1-\beta_4)\beta_1}{G\left(\beta_2 + \beta_1\beta_4 + \beta_1\frac{(1-\beta_4)}{\Phi+1}\right) + (\beta_1+\beta_2)(1-G)(\Phi+1)^{-\mu}}\frac{\Phi}{1+\Phi}. \end{aligned} \tag{151}$$

This expression is increasing in Φ. From (145), Φ moves like w_H/w_L, therefore the labor share decreases in N (the opposite of w_H/w_L) when $H \approx H^P$ (this result may not extend if H^P is far from H when β_4 is close to 1).

Further, we can rewrite (145) as:

$$\frac{w_H H^P}{w_L L} = \frac{\beta_2}{\beta_1} + \frac{(\beta_2 + \beta_1\beta_4)}{\beta_1} \frac{G\Phi\left(\Phi+1\right)^{\mu-1}}{G\left(\Phi+1\right)^{\mu-1} + (1-G)}.$$

We have already derived that an increase in G increases w_H/w_L, therefore, this expression shows that it will increase $\frac{G\Phi(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1}+(1-G)}$. We can then rearrange terms in (151) and write:

$$\begin{aligned} & \frac{\widetilde{r}K + T_m}{w_L L + w_H H} \\ = \; & \frac{\beta_3}{\beta_2+\beta_1} + \frac{(1-\beta_4)\,\beta_1}{\beta_2+\beta_1}\left(\beta_1\beta_4 + \beta_2 + (\beta_2+\beta_1)\,\frac{G\left(\Phi+1\right)^{\mu-1} + (1-G)}{G\Phi\left(1+\Phi\right)^{\mu-1}}\right)^{-1}. \end{aligned}$$

The right hand side is an increasing function of $\frac{G\Phi(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1}+(1-G)}$, which ensures that the labor share decreases in G when $H \approx H^P$. □

7.10.3 Asymptotic behavior

The asymptotic behavior is in line with Proposition 2 but the fact that automation now replaces low-skill workers with a Cobb-Douglas aggregate of capital and high-skill workers limit the ratio between the growth rate of high-skill and low-skill wages. In addition, we here need to consider the long-run behavior of the gross rental rate $\widetilde{r}$. Since r is determined by the Euler equation, then on a path where consumption growth is asymptotically constant, then $\widetilde{r}$ is also asymptotically constant (see (140)). We focus on the case where $G_\infty \in (0,1)$ (although results analogous to those in Proposition 2 could be derived when $G_\infty \in \{0,1\}$) and prove:

Proposition 9. *Consider four processes* $[N_t]_{t=0}^{\infty}$, $[G_t]_{t=0}^{\infty}$, $\left[H_t^P\right]_{t=0}^{\infty}$ *and* $[\widetilde{r}_t]_{t=0}^{\infty}$ *where* $\left(N_t, G_t, H_t^P, \widetilde{r}_t\right) \in (0,\infty) \times [0,1] \times (0,H] \times (0,\infty)$ *for all* t. *Assume that* G_t, g_t^N, H_t^P *and* $\widetilde{r}_t$ *all admit positive and finite limits with* $G_\infty \in (0,1)$. *Then the asymptotic growth rate of high-skill wages* w_{Ht} *and output* Y_t *are*

$$g_\infty^{w_H} = g_\infty^Y = g_\infty^N / \left[(\sigma-1)\left(\beta_2 + \beta_1\beta_4\right)\right], \tag{152}$$

and the asymptotic growth rate of low-skill wages is

$$g_{\infty}^{w_L} = \frac{1+(\sigma-1)\beta_1\beta_4}{1+(\sigma-1)\beta_1} g_{\infty}^{w_H}. \tag{153}$$

Proof. For simplicity we assume that the limits $g_{\infty}^{w_H}, g_{\infty}^{w_L}$ and g_{∞}^{Y} exist (although we could show that formally as we did in Appendix 6.2). Suppose that $g_{\infty}^{w_L} \leq \beta_4 g_{\infty}^{w_H}$. Then Φ_t must either tend toward a positive constant or toward 0, in either case (145) implies that $g_{\infty}^{w_L} = g_{\infty}^{w_H}$, which is a contradiction as $\beta_4 < 1$. Hence it must be that $g_{\infty}^{w_L} > \beta_4 g_{\infty}^{w_H}$, which ensures that $\Phi_t \to \infty$. Using this in (142), we obtain:

$$w_{Ht}^{\beta_2+\beta_1\beta_4} \underset{t\to\infty}{\sim} \frac{\sigma-1}{\sigma} \frac{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}\left(G_{\infty}\varphi\right)^{\frac{1}{\sigma-1}}}{\left(1+\tau_m\right)^{(1-\beta_4)\beta_1}\widetilde{r}_{\infty}^{\beta_3+(1-\beta_4)\beta_1}} N_t^{\frac{1}{\sigma-1}}.$$

This establishes $g_{\infty}^{w_H} = g_{\infty}^{N}/\left[(\sigma-1)\left(\beta_2+\beta_1\beta_4\right)\right]$, from which we can obtain that $g_{\infty}^{Y} = g_{\infty}^{w_H} = g_{\infty}^{N}/\left[(\sigma-1)\left(\beta_2+\beta_1\beta_4\right)\right]$ (using that H_t^P admits a positive limit).

Moreover (145) now implies

$$\begin{aligned} & \frac{w_{Ht}H_{\infty}^P}{w_{Lt}L} \underset{t\to\infty}{\sim} \frac{G_{\infty}\left(\beta_2+\beta_1\beta_4\right)\Phi_t^{\mu}}{\beta_1\left(1-G_{\infty}\right)}, \\ \Longrightarrow \quad & w_{Lt}^{1+\beta_1(\sigma-1)} \underset{t\to\infty}{\sim} \frac{\beta_1\left(1-G_{\infty}\right)\left(1+\tau_m\right)^{(1-\beta_4)\beta_1(\sigma-1)}\widetilde{r}_{\infty}^{(1-\beta_4)\beta_1(\sigma-1)}H_{\infty}^P}{G_{\infty}\left(\beta_2+\beta_1\beta_4\right)\varphi^{\mu}L} w_{Ht}^{1+\beta_4\beta_1(\sigma-1)}, \end{aligned}$$

which implies (153). Since $\frac{1+(\sigma-1)\beta_1\beta_4}{1+(\sigma-1)\beta_1} > \beta_4$, we verify that $g_{\infty}^{w_L} > \beta_4 g_{\infty}^{w_H}$. □

7.10.4 Dynamic equilibrium

We can solve for the dynamic equilibrium as in the baseline model. The long-run elasticity of output with respect to the number of products is now given by $\psi \equiv 1/\left[(\sigma-1)\left(\beta_2+\beta_1\beta_4\right)\right]$. We then introduce the same normalized variables as in the baseline model: $\widehat{V}_t^A$, $\widehat{V}_t^N$, $\widehat{\pi}_t^A$, $\widehat{\pi}_t^N$, $\widehat{h}_t^A$, $\widehat{c}_t$ and $\widehat{v}_t$. We also introduce $\widehat{Y}_t \equiv Y_t N^{-\psi}$ and $\widehat{K}_t \equiv K_t N^{-\psi}$. Finally we now define

$$n_t \equiv N_t^{-\frac{1-\beta_4}{1+\beta_1(\sigma-1)}\frac{\beta_1}{\beta_2+\beta_1\beta_4}}$$

and

$$\omega_t \equiv \left(N_t^{\frac{1-\beta_4}{(\sigma-1)(\beta_2+\beta_1\beta_4)(1+\beta_1(\sigma-1))}} \frac{\widetilde{r}_t^{1-\beta_4} w_{Ht}^{\beta_4}}{w_{Lt}}\right)^{\beta_1(\sigma-1)},$$

so that

$$\left(\frac{w_L}{\widetilde{r}_t^{1-\beta_4} w_H^{\beta_4}}\right)^{\beta_1(1-\sigma)} = \omega_t n_t.$$

The transitional dynamics can then be expressed as a system of differential equations in $\mathbf{x}_t \equiv \left(n_t, G_t, \widehat{K}_t, \widehat{h}_t^A, \widehat{v}_t, \widehat{c}_t\right)$ where the first three variables are state variables and the last three control variables.

Equation (39) still applies, therefore, we get using (141) that

$$\pi_t^A = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\left(\varphi\left(\left((1+\tau_m)\,\widetilde{r}_t\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu}\left(\frac{w_H^{\beta_2}\widetilde{r}^{\beta_3}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}\right)^{1-\sigma} Y_t.$$

We can rewrite this as

$$\widehat{\pi}_t^A = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\left(\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}\right)^{\sigma-1}\left(\widetilde{r}_t^{(1-\beta_4)\beta_1+\beta_3}\widehat{v}_t^{\beta_2+\beta_4\beta_1}\right)^{1-\sigma}\left(\varphi\,(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu}\widehat{Y}_t. \tag{154}$$

We can derive π_t^N similarly and we find

$$\widehat{\pi}_t^N = \omega_t n_t\left(\varphi\,(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu}\widehat{\pi}_t^A \tag{155}$$

(30) is now replaced by

$$\dot{n}_t = -\frac{1-\beta_4}{1+\beta_1(\sigma-1)}\frac{\beta_1}{\beta_2+\beta_1\beta_4}g_t^N n_t. \tag{156}$$

(31) still applies and so does (32). Because of the automation tax (33) is replaced by

$$\left(r_t - (\psi-1)\,g_t^N\right)\widehat{V}_t^N = \widehat{\pi}_t^N + \eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa}\left(\widehat{V}_t^A - \widehat{V}_t^N\right) - (1+\tau_a)\,\widehat{v}_t\widehat{h}_t + \dot{\widehat{V_t^N}} \tag{157}$$

and (34) by

$$\kappa\eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa-1}\left(\widehat{V}_t^A - \widehat{V}_t^N\right) = (1+\tau_a)\,\widehat{v}_t. \tag{158}$$

Combining (155), (157), (158) and (17), we now obtain:

$$\dot{\widehat{v}}_t = \widehat{v}_t\left(\widetilde{r}_t - \Delta - (\psi-1)\,g_t^N - \gamma\omega_t n_t\left(\varphi\,(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu}\frac{\widehat{\pi}_t^A}{\widehat{v}_t} - \gamma\,(1+\tau_a)\,\frac{1-\kappa}{\kappa}\widehat{h}_t^A\right). \tag{159}$$

Following the same steps as those used to derive (37), we now obtain:

$$\begin{aligned}
&\dot{\widehat{h}}_t^A \\
= \;& \gamma \frac{\widehat{h}_t^A}{1-\kappa}\left(\omega_t n_t\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)}+\left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{-\mu}\frac{\widehat{\pi}_t^A}{\widehat{v}_t}+\left(1+\tau_a\right)\frac{1-\kappa}{\kappa}\widehat{h}_t^A\right) \\
& -\frac{\kappa\eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa}}{(1-\kappa)\left(1+\tau_a\right)}\left(1-\omega_t n_t\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)}+\left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{-\mu}\right)\frac{\widehat{\pi}_t^A}{\widehat{v}_t}+\eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa+1} \\
& +\frac{\widetilde{\kappa}}{1-\kappa}\left(\eta\left(\widehat{h}_t^A\right)^{\kappa+1}G_t^{\widetilde{\kappa}-1}\left(1-G_t\right)-g_t^N\widehat{h}_t^A\right)
\end{aligned} \tag{160}$$

Further, (18) still applies and we can rewrite it as:

$$\dot{\widehat{c}}_t=\frac{\widehat{c}_t}{\theta}\left(\widetilde{r}_t-\left(\rho+\Delta+\theta\psi g_t^N\right)\right). \tag{161}$$

Finally, we can rewrite (23) as

$$\dot{\widehat{K}}_t=\widehat{Y}_t-\widehat{c}_t-\left(\Delta+\psi g_t^N\right)\widehat{K}_t \tag{162}$$

Equations (156), (31), (159), (160), (161) and (162) form a system of differential equations which depend on $\widehat{Y}_t$, $\widehat{\pi}_t^A$, $\widetilde{r}_t$ and g_t^N.

(142) implies

$$\frac{\sigma}{\sigma-1}\frac{\widehat{v}^{\beta_2+\beta_4\beta_1}\widetilde{r}^{\beta_3+\beta_1(1-\beta_4)}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}\left(G\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)}+\left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu}+\left(1-G\right)\omega_t n_t\right)^{\frac{1}{1-\sigma}}=1,$$

so that

$$\widetilde{r}=\left[\frac{\sigma-1}{\sigma}\frac{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}{\widehat{v}^{\beta_2+\beta_4\beta_1}}\left(G\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)}+\left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu}+\left(1-G\right)\omega_t n_t\right)^{\frac{1}{\sigma-1}}\right]^{\frac{1}{\beta_3+\beta_1(1-\beta_4)}}, \tag{163}$$

which defines $\widetilde{r}$ as a function of $\mathbf{x}_t$ and ω_t. (144) can be written as:

$$H_t^P=\frac{\widetilde{r}_t\widehat{K}_t}{\widehat{v}_t}\frac{G_t\left(\beta_2+\frac{\beta_1\beta_4\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}{(\omega_t n_t)^{\frac{1}{\mu}}+(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}\right)\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)}+\left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu}+\beta_2\left(1-G_t\right)\omega_t n_t}{G_t\left(\beta_3+\frac{\beta_1(1-\beta_4)\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)-1}}{(\omega_t n_t)^{\frac{1}{\mu}}+\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}\right)\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)}+\left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu}+\beta_3\left(1-G_t\right)\omega_t n_t}, \tag{164}$$

which gives, together with (163), H_t^P as a function of $\mathbf{x}_t$ and ω_t. g_t^N still obeys (46), which then defines it as a function of $\mathbf{x}_t$ and ω_t.

Combine (148), (143), (144) and (149) to obtain:

$$\frac{Y}{w_H H^P} = \frac{\frac{\sigma}{\sigma-1}\left(G\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} + \left(1-G\right) w_L^{\beta_1(1-\sigma)}\right)}{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi\left(\left((1+\tau_m)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi\left(\left((1+\tau_m)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}\right)\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} + \beta_2\left(1-G\right) w_L^{\beta_1(1-\sigma)}},$$

which we can rewrite as

$$\widehat{Y}_t = \frac{\sigma}{\sigma-1}\frac{\left[G_t\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + \left(1-G\right)\omega_t n_t\right]\widehat{v}_t H_t^P}{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}{\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}+(\omega_t n_t)^{\frac{1}{\mu}}}\right)\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + \beta_2\left(1-G\right)\omega_t n_t}. \tag{165}$$

This expression, with the previous equations, gives $\widehat{Y}_t$ as a function of $\mathbf{x}_t$ and ω_t. (154) then ensures that $\widehat{\pi}_t^A$ is defined as a function $\mathbf{x}_t$ and ω_t.

Finally, from (143) we obtain:

$$\omega_t = \left[\frac{\left(\frac{\widehat{v}_t}{\widetilde{r}_t}\right)^{1-\beta_4}\frac{H_t^P}{L}\beta_1\left(G\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)}\left(\omega_t n_t\right)^{-\frac{1}{\mu}}+1\right)^{\mu-1} + \left(1-G\right)\right)}{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}{(\omega_t n_t)^{\frac{1}{\mu}}+\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}\right)\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + \beta_2\left(1-G\right)\omega_t n_t}\right]^{\frac{\beta_1(1-\sigma)}{1+\beta_1(\sigma-1)}}, \tag{166}$$

which implicitly defines ω_t as a function of $\mathbf{x}_t$. Hence, together with (163), (164), (46), (165), (154) and (166), the system formed by (156), (31), (159), (160), (161) and (162) describes the dynamic equilibrium. We then obtain

Proposition 10. *Assume that*

$$\rho\left(\frac{\left(1+\tau_a\right)^{\kappa}}{\kappa^{\kappa}\left(1-\kappa\right)^{1-\kappa}\eta}\left(\frac{\rho}{\gamma}\right)^{1-\kappa} + \frac{1}{\gamma}\right) < \psi H \tag{167}$$

is satisfied, then the economy admits a steady-state $\left(n^*, G^*, \widehat{K}^*, \widehat{h}^{A*}, \widehat{v}^*, \widehat{c}^*\right)$ *with* $n^* = 0$, $G^* \in (0,1)$ *and* $g^{N*} > 0$. g^{N*}, G^* *and* $\widehat{h}^{A*}$ *are independent of* τ_m.

Proof. As before, we directly get that in a steady-state with $g^{N*} > 0$, we must have $n^* = 0$. (166) then implies that ω^* is a constant defined by

$$\omega^* = \left[\left(\frac{\widehat{v}^*}{\widetilde{r}^*} \right)^{1-\beta_4} \frac{H^{P*}}{L} \frac{\beta_1 (1 - G^*) (1 + \tau_m)^{(1-\beta_4)(\sigma-1)}}{G^* (\beta_2 + \beta_1 \beta_4) \varphi^\mu} \right]^{\frac{\beta_1(1-\sigma)}{1+\beta_1(\sigma-1)}}.$$

This guarantees that in such a steady-state, $w_{Lt} \sim \omega^{*\frac{1}{\beta_1(1-\sigma)}} \widetilde{r}^{*1-\beta_4} \widehat{v}^{*\beta_4} N_t^{\frac{1+\beta_4\beta_1(\sigma-1)}{1+\beta_1(\sigma-1)} \frac{1}{(\sigma-1)(\beta_2+\beta_1\beta_4)}}$ such that $g_\infty^{w_L} = \frac{1+\beta_4\beta_1(\sigma-1)}{1+\beta_1(\sigma-1)} g_\infty^{w_H}$ as stipulated in Proposition 9.

In addition, (161) implies that in steady-state,

$$\widetilde{r}^* = \rho + \Delta + \theta\psi g^{N*}. \tag{168}$$

(164) implies that

$$H^{P*} = \frac{\widetilde{r}^* \widehat{K}^*}{\widehat{v}^*} \frac{(\beta_2 + \beta_1\beta_4)}{(\beta_3 + \beta_1 (1 - \beta_4))}. \tag{169}$$

Then (165) implies that

$$\widehat{Y}^* = \frac{\sigma}{(\sigma - 1)(\beta_2 + \beta_1\beta_4)} \widehat{v}^* H^{P*}. \tag{170}$$

We then get that (154) implies that

$$\frac{\widehat{\pi}^{A*}}{\widehat{v}^*} = \frac{(\sigma-1)^{\sigma-2}}{\sigma^{\sigma-1}} \frac{\left(\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}\right)^{\sigma-1}}{(\beta_2 + \beta_1\beta_4)} \left(\widetilde{r}^{*(1-\beta_4)\beta_1+\beta_3} \widehat{v}^{*\beta_2+\beta_4\beta_1} (1 + \tau_m)^{(1-\beta_4)\beta_1} \right)^{1-\sigma} \varphi^\mu H^{P*}. \tag{171}$$

(163) gives

$$\widetilde{r}^* = \left[\frac{\sigma - 1}{\sigma} \frac{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}{\widehat{v}^{*\beta_2+\beta_4\beta_1}} \frac{(G^*\varphi^\mu)^{\frac{1}{\sigma-1}}}{(1 + \tau_m)^{(1-\beta_4)\beta_1}} \right]^{\frac{1}{\beta_3+\beta_1(1-\beta_4)}}. \tag{172}$$

Therefore (171) simplifies into

$$\frac{\widehat{\pi}^{A*}}{\widehat{v}^*} = \psi \frac{H^{P*}}{G^*}, \tag{173}$$

just as in the baseline model. Then (159) and (168) together imply that

$$\widehat{h}^{A*} = \frac{\kappa}{\gamma (1 + \tau_a)(1 - \kappa)} \left(\rho + ((\theta - 1)\psi + 1) g^{N*} \right). \tag{174}$$

This defines $\widehat{h}^{A*}$ as an increasing function of g^{N*}. Further, in steady-state G^* still obeys

(49) and H^{P*} obeys (51), which imply that G^* and H^{P*} also be defined as function of g^{N*}.

(173), (160), (49), (174) then lead to

$$\frac{1-\kappa}{\kappa}\frac{\gamma G^*\left(1+\tau_a\right)\widehat{h}^{A*}}{\psi H^{P*}}\left(\frac{1+\tau_a}{\kappa\eta G_t^{\widetilde{\kappa}}}\left(\widehat{h}^{A*}\right)^{1-\kappa}+\frac{1}{\gamma}\right)=1, \tag{175}$$

which up to the term $1+\tau_a$ is the same as (53) in the baseline case. Therefore following the same reasoning, there exists a steady-state with $g^{N*}>0$ and $G^*\in(0,1)$ as long as (167) is satisfied. As (175), (49), (51) and (174) are independent of τ_m, so are g^{N*}, $\widehat{h}^{A*}$ (now given by (174)), G^* (given by (49)) and H^{P*} (given by (51)).

We further obtain $\widetilde{r}^*$ through (168), which must be independent of τ_m as well. We then get $\widehat{v}^*$ through (172) as

$$\widehat{v}^*=\left[\frac{\sigma-1}{\sigma}\frac{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}{\widetilde{r}^{*\beta_3+\beta_1(1-\beta_4)}}\frac{\left(G^*\varphi^{\mu}\right)^{\frac{1}{\sigma-1}}}{\left(1+\tau_m\right)^{(1-\beta_4)\beta_1}}\right]^{\frac{1}{\beta_2+\beta_4\beta_1}}.$$

We then get $\widehat{K}^*$ through (169) and $\widehat{c}^*$ from (162) which, using (170), implies:

$$\widehat{c}^*=\frac{\sigma}{(\sigma-1)\left(\beta_2+\beta_1\beta_4\right)}\widehat{v}^*H^{P*}-\left(\Delta+\psi g^{N^*}\right)\widehat{K}^*.$$

□

Further if $\tau_a=\tau_m=0$, $g^{N*},G^*,\widehat{h}^{A*}$ are determined by the same equations are in the baseline model except that the definition of ψ has changed. It is then direct that Proposition 7 extends to this case.

7.10.5 Short-run effect of a machine tax

We look at the short-run effect of a machine tax on wages, taking as given the allocation of high-skill labor between innovation and production and the total supply of capital (but not its allocation or the rental rate). Using (146), we can rewrite (142) and (144) as:

$$N^{\frac{1}{1-\sigma}}\frac{\sigma}{\sigma-1}\frac{w_H^{\beta_2}\widetilde{r}^{\beta_3}w_L^{\beta_1}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}\left(G\left(\Phi+1\right)^{\mu}+1-G\right)^{\frac{1}{1-\sigma}}=1, \tag{176}$$

$$\frac{\widetilde{r}K}{w_HH^P}=\frac{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)\left(\Phi+1\right)^{\mu}+\beta_3\left(1-G\right)}{G\left(\beta_2+\frac{\beta_1\beta_4\Phi}{1+\Phi}\right)\left(\Phi+1\right)^{\mu}+\beta_2\left(1-G\right)}. \tag{177}$$

Then, totally log-differentiate (146), (145) and (177) to get:

$$\frac{1}{\varepsilon-1}\widehat{\Phi} = \widehat{w}_L - \widehat{w}_H + (1-\beta_4)\left(\widehat{w}_H - \widehat{\widetilde{r}}\right) - (1-\beta_4)\,\widehat{1+\tau_m} \tag{178}$$

$$\widehat{w}_H - \widehat{w}_L = \left(\begin{array}{c} \frac{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}}{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}+\beta_2(1-G)}\left(\mu+\frac{\beta_1\beta_4\frac{1}{1+\Phi}}{\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}}\right) \\ +\frac{(1-\mu)G(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1}+(1-G)} \end{array}\right)\frac{\Phi\widehat{\Phi}}{\Phi+1}, \tag{179}$$

$$\begin{aligned} \widehat{\widetilde{r}} - \widehat{w}_H \;=\; & \frac{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}}{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}+\beta_3(1-G)}\left(\frac{\mu\Phi}{\Phi+1}\widehat{\Phi}+\frac{\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\left(\frac{\widehat{\Phi}}{1+\Phi}-\widehat{1+\tau_m}\right)}{\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}}\right) \\ & -\frac{G\left(\beta_2+\frac{\beta_1\beta_4\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}}{G\left(\beta_2+\frac{\beta_1\beta_4\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}+\beta_2(1-G)}\left(\mu+\frac{\beta_1\beta_4\frac{1}{1+\Phi}}{\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}}\right)\frac{\Phi\widehat{\Phi}}{1+\Phi} \end{aligned} \tag{180}$$

Combine (178), (179) and (180) to get:

$$\begin{aligned} & \left(\begin{array}{c} \frac{1}{\varepsilon-1}\frac{1+\Phi}{\Phi}+\frac{\beta_4 G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}}{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}+\beta_2(1-G)}\left(\mu+\frac{\beta_1\beta_4\frac{1}{1+\Phi}}{\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}}\right) \\ +\frac{(1-\mu)G(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1}+(1-G)}+\frac{(1-\beta_4)G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}}{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}+\beta_3(1-G)}\left(\mu+\frac{\frac{\beta_1(1-\beta_4)}{(1+\tau_m)(1+\Phi)}}{\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}}\right) \end{array}\right)\frac{\Phi\widehat{\Phi}}{1+\Phi} \\ = \; & -(1-\beta_4)\,\beta_3\frac{G(\Phi+1)^{\mu}+(1-G)}{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}+\beta_3(1-G)}\widehat{1+\tau_m}. \end{aligned} \tag{181}$$

As $\mu < 1$, the coefficient in front of $\widehat{\Phi}$ on the LHS is positive, so that Φ is decreasing in the machine tax τ_m. Using (179), we also get that the skill premium decreases in the machine tax.

Totally log-differentiating (176), one gets

$$\widehat{w}_L = \frac{\mu}{\sigma-1}\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+1-G}\frac{\Phi\widehat{\Phi}}{\Phi+1} - (1-\beta_1)\left(\widehat{w}_H-\widehat{w}_L\right) - \beta_3\left(\widehat{\widetilde{r}}-\widehat{w}_H\right).$$

Plugging (179) and (180) and (181), we get:

$$\widehat{w}_L = -\left(\begin{array}{c} \frac{(1-\mu)G(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1}+(1-G)}\left[1-\beta_1+\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+(1-G)}\frac{\beta_1\Phi}{(1+\tau_m)(1+\Phi)}\right]-\frac{\beta_1}{\varepsilon-1}\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+1-G}\frac{\Phi\widehat{\Phi}}{\Phi+1}\frac{\tau_m}{1+\tau_m} \\ +\frac{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}}{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}+\beta_2(1-G)}\left(\mu+\frac{\beta_1\beta_4\frac{1}{1+\Phi}}{\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}}\right)\left[\beta_2+\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+(1-G)}\frac{\beta_4\beta_1\Phi}{(1+\tau_m)(1+\Phi)}\right] \\ +\frac{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}}{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}+\beta_3(1-G)}\left(\mu+\frac{\frac{\beta_1(1-\beta_4)}{(1+\tau_m)(1+\Phi)}}{\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}}\right)\left[\beta_3+\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+(1-G)}\frac{(1-\beta_4)\beta_1\Phi}{(1+\tau_m)(1+\Phi)}\right] \end{array}\right)\frac{\Phi\widehat{\Phi}}{\Phi+1}$$

For a small tax on machines (i.e. around $\tau_m = 0$), the only negative term inside the parenthesis drops out, so that the introduction of a small tax on machines leads to an increase in low-skill wages. Therefore we have established:

Proposition 11. *On impact, a tax on machines reduces the skill premium and a small tax on machines increases low-skill wages.*

7.11 Appendix to the Quantitative Exercise

Section 7.11.1 describes the calibration technique and Section 7.11.2 the data used for the calibration and the patent data presented in Figure 1.C and 7. Section 7.11.3 discusses how parameters are identified. Section 7.11.4 presents details on the constant G alternative model discussed in Section 4.2. Section 7.11.5 shows that the data still require an increase in the share parameter G when there is constant labor-augmenting technical change. Finally, Section 7.11.6 carries an analysis of the effect of automation taxes in our calibrated model.

7.11.1 Technique

We choose parameters to minimize the log-deviation of predicted and observed variables for the four time paths of the skill-premium, the labor-share of GDP, stock of equipment over GDP and an index of GDP per hours worked. That is, for a given set of parameters b the model produces predicted output of $\hat{Y}_i = \left\{\hat{Y}_{i,t}\right\}_{t=1}^{T_i}$ for each of these four paths from 1963 and until 2007 for the labor share, skill-premium, and GDP per hour, and 2000 for equipment over GDP (due to data limitations from Cummins and Violante, 2002). We let $\hat{Y}(b) = \{\hat{Y}_i(b)\}_{i=1}^{4}$ as the combined vector of these paths and make explicit the dependency on the parameters b. Y is the corresponding vector of actual values. We

then solve :

$$min_b(log(\hat{Y}(b)) - log(Y))'W(log(\hat{Y}(b)) - log(Y)),$$

where W is a diagonal matrix of weights. In a previous version of the paper (Hémous and Olsen, 2016) we articulated a stochastic version of our model by introducing autocorrelated measurement errors. Here we choose a much simpler approach and simply choose "reasonable" weights based on how easily the model matches the path. In particular, the diagonal elements are 4 for the skill-premium, though 10 for the first 5 years, 10 for the labor share, 1 for GDP/hours and 2 for equipment over GDP. For a given starting value of b we then run 12 estimations based on "nearby" randomly chosen parameters. We choose the best fit of these 13 (12 plus the original starting point), take that value as the next starting value and repeat the step. We continue this process until 100 steps (1200 nearby simulations) have not improved the fit. We do this for 10 (substantially) different starting points. They all give the same result. There is little substantial difference between the Bayesian approach taken previously and the one pursued here.

7.11.2 Data

Calibration Data. We do not seek to match the skill-ratio H/L but take it as exogenously given. We normalize $H + L = 1$, throughout. The skill-ratio is taken from Acemoglu and Autor (2011). However, since our estimation requires a skill-ratio both before and after the period 1963-2007 we match the observed path of the log of the skill-ratio to a "generalized" logistical function of the form:

$$\frac{\alpha}{1 + exp\left(\frac{\mu - t}{s}\right)} + \beta,$$

where (α, β, μ, s) are parameters to be estimated. We use the observed skill-ratio in the period $1963 - 2007$ and the predicted values outside of this time interval. Yet, the fit is so good that there is no visual difference in the match of the four time periods between this approach and using the predicted value in the interval $1963 - 2007$. The skill-premium is taken from Autor (2014) which extends the data of the Acemoglu and Autor (2011) until 2012. The labor share is the BLS's labor share in the non-farm business sector. We take GDP per hours worked from the series on non-farm business from the BLS (series PRS85006092). Capital equipment is calculated as follows: We follow KORV and use quality-adjusted price indices of equipment from Cummins and Violante (2002) who update the series from Gordon (1990). We combine two different series. First, we

use NIPA data on private investment in equipment excluding transportation equipment (Tables 1.5 and 5.3.5. from NIPA). We iteratively construct an index for the stock of private real capital equipment by assuming a depreciation rate of 12.5 per cent (as Krusell et al., 2000) and using the price index for private equipment from Cummins and Violante (2002). We start this approach in 1947 but only use the stock from 1963 onwards. We combine this with the growth rate of real private GDP to get an index for equipment over GDP. We match this index to the NIPA private equipment capital stock (excluding transportation) over (private) GDP number for 1963 to get a series in absolute value. To this, we add software, but following the suggestion of Cummins and Violante (2002), we use the NIPA data on the stock of software over GDP (table 2.1 from NIPA). We add these two values to get our combined stock of equipment (+software) over GDP.

Labor costs data. We combine three sources of data. First, two indices from FRED (Federal Reserve Bank of St. Louis), the *Average Hourly Earnings of Production and Nonsupervisory Employees, Total Private (AHETPI_NBD20120101)* which contains the wages and salaries of workers and Nonfarm Business Sector and *Compensation Per Hour*, which contains total compensation of workers. These indices give the trend growth. To couple this with a level, we use the BLS Bureau of Labor Statistics from March 1991 which gives wages and salaries as a share of total compensation of 72.3 per cent for private workers in the United States for 1991. We combine these to gives a time path since 1964 of share of total compensation going to wages and salaries (we equate 1963 to 1964). This path is consistent with other more recent time trends for the BLS. For instance, the December 2019 release of the Employment Cost Index from the BLS shows a small decline in the share of total compensation going to wages of .7 per cent from 2005 to 2019 (combining tables 4 and 8) which is consistent with our measure which shows a small increase of 1 per cent over the same time period. Finally, though the share of total compensation not going to wages and salaries is not constant across occupational groups it is substantial for all groups and is never below 27 per cent for full-time workers (https://www.bls.gov/news.release/ecec.t05.htm).

Patent data on automation innovations. The data for Figure 7.A are taken from Mann and Püttmann (2018) who classify USPTO patents granted from 1976. Given that they classify patents according to their grant years, we lag all their numbers by 2 years to reproduce an approximate time lag of 2 years between application year (which is closer to the year of innovation) and grant year. They find that commuting zones exposed

to industries with a higher level of automation experience a decline in manufacturing employment (and an increase in overall employment).

Dechezleprêtre et al. (2019) classify patents in machinery as automation versus non-automation. Their classification method follows two steps. First, they classify technological codes (IPC and CPC codes, mostly at the 6 digit level) by computing the frequency of certain keywords which have been related to automation (such as "robot", "automation", "computer numerical control", etc...) for each technological code in machinery. They identify automation technological codes as those with a high share of patents with a keyword (in the top 5 percent of the distribution) and non-automation codes as those with a low share (in the bottom 60 percent). Second, they define automation patent as patents with at least one automation technological code and non-automation patents as patents with only non-automation codes. They show that the share of automation patents in machinery in a sector is correlated with a decline in routine manual and cognitive tasks, and an increase in the high-skill to low-skill employment ratio. Then, they use cross-country variation in wages and variations in firms' exposures to different countries, to show in firm-level regressions, that an increase in low-skill wages leads to an increase in automation innovations but not in non-automation innovations. They classify patents using the PATSTAT database which starts before 1976. In Figure 7.B, we use granted patents at the USPTO.

7.11.3 Parameters identification

In this section, we discuss how our parameters are identified, first by carrying a back-of-the-envelope calibration, second by computing the elasticities of the initial and final values of the series we match with respect to the parameters, and third by computing how precisely each parameter is identified. We then discuss specifically how $\tilde{\kappa}$ is determined and finally carry an out-of-sample prediction exercise, where we only use the first 30 years of the data to calibrate our parameters.

Back-of-the-envelope calibration. We first study how the production parameters σ, β_1, β_2, β_4 and Δ would be identified under a naive back-of-the-envelope calibration, where we assume that in 1963 the U.S. economy was in the first phase while in 2012, it was in the third phase. Since both assumptions are actually not met in our estimation, this naive calibration gives parameters that are still far from those which we actually estimate. Nevertheless, the exercise is informative to understand how these production parameters are related to moments in the data.

Assuming that the economy in 1963 is close to the first phase, and using (143), we get that the skill premium must obey:

$$\frac{w_{H1963}}{w_{L1963}} \approx \frac{\beta_2}{\beta_1} \frac{L_{1963}}{H_{1963} - \frac{1}{\gamma} g^N_{1963}}.$$

Further, using that most high-skill workers work in production, such that $\frac{1}{\gamma} g^N_{1963}$ is small relative to H, we obtain

$$\frac{\beta_2}{\beta_1} \approx \frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}}, \tag{182}$$

so that the ratio β_2/β_1 is determined by the ratio between the high-skill wage bill and the low-skill wage bill. Because the economy is in fact not in the first phase in 1963 (with an equipment stock to GDP ratio which is not 0), this approximation is likely to overstate the ratio β_2/β_1. Similarly, using (143), (144), and (150), we get that the labor share in 1963 should obey

$$ls_{1963} \approx \frac{\beta_2 \frac{H}{H - \frac{1}{\gamma} g^N_{1963}} + \beta_1}{\frac{\sigma}{\sigma - 1} + \beta_2 \frac{\frac{1}{\gamma} g^N_{1963}}{H - \frac{1}{\gamma} g^N_{1963}}},$$

which simplifies into

$$ls_{1963} \approx \frac{\sigma - 1}{\sigma} \left(\beta_2 + \beta_1 \right), \tag{183}$$

if most high-skill workers are in production. Therefore, given σ, the initial labor share determines β_3, the 'external' capital share. We can then combine (182) and (183) to obtain

$$\beta_1 \approx \frac{1}{\frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}} + 1} \frac{ls_{1963}}{1 - \frac{1}{\sigma}}, \tag{184}$$

so that β_1 which is the Cobb-Douglas share for low-skill workers in Phase 1 is given by the labor share in 1963 and the ratio between the high-skill wage bill and the low-skill wage bill, and σ which determines mark-ups.

Combining (153) and (153), we get that if the economy is close to its asymptotic steady-state in 2012, the growth rate of the skill premium is given by

$$g^{sp}_{2012} \approx \frac{\beta_1 (\sigma - 1)(1 - \beta_4)}{1 + \beta_1 (\sigma - 1)} g^{GDP}_{2012}. \tag{185}$$

Using (143), (144) (150), the labor share now obeys:

$$ls_{2012} \approx H\left[\frac{\sigma H}{(\sigma-1)(\beta_2+\beta_1\beta_4)} - \left(\frac{\sigma}{(\sigma-1)(\beta_2+\beta_1\beta_4)} - 1\right)\left(\frac{1}{\gamma}g^N_{1963} + H^A_{1963}\right)\right]^{-1},$$

which under the assumption that most high-skill workers are in production would simplify again into

$$ls_{2012} \approx \frac{\sigma-1}{\sigma}(\beta_2+\beta_1\beta_4). \tag{186}$$

Combining (182), (183) and (186) we obtain:

$$\beta_4 \approx 1 - \left(1 - \frac{ls_{2012}}{ls_{1963}}\right)\left(\frac{w_{H1963}H_{1963}}{w_{L1963}L_{1963}} + 1\right). \tag{187}$$

Therefore, in this approximation, β_4 is identified through the decline in the labor share and the initial wage bill ratio between high-skill and low-skill workers. In the data the labor share does not monotonically decline. To understand how the parameters are identified, we replace ls_{2012} by the lowest value over 1963-2012 (which is 57%) and ls_{1963} by the highest value (64.6%). With $\frac{w_{H1963}H_{1963}}{w_{L1963}L_{1963}} = 0.576$, we then obtain $\beta_4 \approx 0.82$. This is higher than the value we actually end up finding ($\beta_4 = 0.73$), mostly because the economy is still far from its steady-state in 2012 (so that ls_{2012} is higher than the asymptotic value of the labor share).

Using (184), (185) and (187) we obtain:

$$\sigma \approx \frac{1}{ls_{1963}\left[\frac{g^{GDP}_{2012}}{g^{sp}_{2012}}\left(1 - \frac{ls_{2012}}{ls_{1963}}\right) - \frac{1}{\frac{w_{H1963}H_{1963}}{w_{L1963}L_{1963}}+1}\right]}.$$

that is given the initial wage bill ratio and the labor shares in 1963 and 2012, which inform us about β_1, β_2 and β_4, σ is determined by the ratio between the growth rate of GDP and that of the skill-premium in the third phase. The larger is σ, the more automated firms gain over non-automated ones and therefore the more the skill premium rises relative to GDP: hence a lower $\frac{g^{GDP}_{2012}}{g^{sp}_{2012}}$ is associated with a larger σ. When using the last 10 years to determine $\frac{g^{GDP}_{2012}}{g^{sp}_{2012}}$, we find that $\sigma \approx 6.77$, while our estimation procedure leads to $\sigma = 5.94$.

Given σ, one can then find β_1 using (184), we find $\beta_1 \approx 0.48$, below but not too far from the estimated value of 0.59 (this approximation is not too sensitive on σ provided that σ is large enough). Using (182), we then obtain $\beta_2 \approx 0.28$ which is higher than the

estimated value of 0.18, in line with the fact that (182) gives an overestimate of β_2/β_1.

To get a proxy for Δ, we look at the steady-state value for the equipment to GDP ratio. Using (169), (170) and the definition of GDP, we obtain that

$$\frac{\widehat{K}^*}{\widehat{GDP}^*} = \frac{1}{\widetilde{r}^*} \frac{\beta_3 + \beta_1 (1-\beta_4)}{\frac{\sigma}{\sigma-1} + (\beta_2 + \beta_1\beta_4)\left(\frac{H}{H^{P*}} - 1\right)}.$$

Denote by K_{eq} the stock of capital used as equipment, we get

$$\widehat{K}^*_{eq} = \frac{\beta_1 (1-\beta_4)}{\beta_3 + \beta_1 (1-\beta_4)} \widehat{K}^*,$$

since in steady-state the economy is Cobb-Douglas with a total physical capital share of $\beta_3 + \beta_1 (1-\beta_4)$ and an equipment of share $\beta_1 (1-\beta_4)$. Using (168), we obtain

$$\frac{\widehat{K}^*_{eq}}{\widehat{GDP}^*} = \frac{1}{\rho + \Delta + \theta g^{GDP*}} \frac{\beta_1 (1-\beta_4)}{\frac{\sigma}{\sigma-1} + (\beta_2 + \beta_1\beta_4)\left(\frac{H}{H^{P*}} - 1\right)}.$$

Therefore, assuming that in 2000 (the last year for which we have data on the equipment to GDP ratio), we are close to the steady-state, and that most high-skill workers are in production, we get

$$\rho + \Delta + \theta g^{GDP}_{2000} = \left(\frac{\sigma-1}{\sigma}\right) \frac{\beta_1 (1-\beta_4)}{\frac{K_{eq,2000}}{GDP_{2000}}} \approx 0.051 \tag{188}$$

using the values computed above. It is therefore not surprising that we find a low Δ in the estimation. This is due in particular to the high-level of $\frac{K_{eq,2000}}{GDP_{2000}} = 1.5$ (with the actual estimated values for σ, β_1 and β_4 we would still find that $\rho + \Delta + \theta g^{GDP}_{2000} \approx 0.088$).

Parameters' effect on empirical moments. We now show the role that parameters have on the empirical moments which allows us to identify what features of the data pin down the parameters. Taking as our starting point the parameter estimates of Section 4, we iteratively change each one by 2% and show the resulting effects on the initial (1963) and the final (2012) values of each of the four empirical paths. Table 6 reports the elasticities (note that β_3 is completely determined by β_1 and β_2).

The initial skill premium is most strongly affected by the production function parameters β_1, β_2 and β_4: A higher share of high-skill workers in production, β_2, directly increases the skill-premium. A higher value of β_4 makes automation more expensive, which increases the demand for low-skill workers and reduces the skill premium. A

higher β_1 implies a lower β_3 which reduces the role of structural capital. This reduces the rental rate of capital, which increases the use of capital equipment and thereby the skill-premium. β_2 has the opposite effect on the skill premium in 2012. A higher β_2 reduces the multiplier of N_t on output, Y_t which reduces the growth rate of the economy. The automation technology parameters, κ, $\tilde{\kappa}$, η also have a large effect on the skill premium in 2012.

The initial labor share depends on β_1, β_2 and β_4, the latter having a much larger effect in 2012 since the share of automated products is much larger.

$GDP/labor$ is mechanically affected negatively by higher σ since we keep the stock of products in 1963 constant. Both β_1 and β_2 reduce the importance of structural capital and thereby have a negative effect on $GDP/labor$ in 1963 as the stock of capital is sufficiently large. In 2012 $\sigma, \beta_{1,}\beta_2, \beta_4$ all reduce the multiplier of N_t on Y_t and therefore $GDP/labor$. The innovation parameters γ, η lead to higher growth and therefore higher $GDP/labor$ in 2012, though naturally not in 1963.

Capital equipment / GDP in 1963 depends positively on β_1 and negatively on β_4 because the initial capital stock is fixed. For 2012, a higher β_4 increases the cost of automation and thereby reduces K_{eq}/GDP. Horizontal innovation productivity, γ, encourages more innovation. This drives up the wage of high-skill workers in 1963, makes automation more expensive and reduces K_{eq}/GDP. It further increases the growth rate of the economy and reduces G_{2012} such that K_{eq}/GDP in 2012 is also lower. Finally, higher productivity of machines, $\tilde{\varphi}$, shifts capital into equipment and consequently raises K_{eq}/GDP.

Precision of the parameters. In the following, we calculate the effect the parameters have on the aggregate final moment. We do this allowing for all the other parameters to adjust, illustrating how precisely each of the parameters are determined. Since deviations from the minimum parameter values are naturally second order we do not compute elasticities. Instead, for a given parameter θ_i consider

$$V(\theta_i, \bar{\theta}_{-i}(\theta_i)),$$

where $\bar{\theta}_{-i}(\theta_i)$ are the parameters that minimize V for any given θ_i and $\bar{\theta}_i = argmin_{\theta_i} V(\theta_i, \bar{\theta}_{-i}(\theta_i))$ is the minimizing value of θ_i. Consequently, a Taylor expansion around $\bar{\theta}_i$ yields:

$$\frac{V(\theta_i, \bar{\theta}_{-i}(\theta_i)) - V(\bar{\theta}_i, \bar{\theta}_{-i}(\bar{\theta}_i))}{V(\bar{\theta}_i, \bar{\theta}_{-i}(\bar{\theta}_i))} / \left(\frac{\theta_i - \bar{\theta}_i}{\bar{\theta}_i}\right)^2 \approx \frac{1}{2} \frac{1}{V(\bar{\theta}_i, \bar{\theta}_{-i}(\bar{\theta}_i))} \frac{d^2 V(\bar{\theta}_i, \bar{\theta}_{-i}(\bar{\theta}_i))}{d\theta_i^2} \bar{\theta}_i^2.$$

Parameters	Skill premium		Labor share		GDP/labor		K_{eq}/GDP	
	1963	2012	1963	2012	1963	2012	1963	2012
σ	−0.1	0.0	0.1	0.0	−0.5	−1.9	2.1	1.1
ϵ	−0.1	−0.1	0.0	0.0	0.0	−0.2	−0.3	−0.2
β_1	0.3	−0.1	0.7	0.6	−1.0	−2.6	7.7	1.8
γ	0.1	−0.1	0.0	0.1	0.0	0.7	−0.2	−0.5
$\tilde{\kappa}$	−0.2	−0.2	0.0	0.1	0.0	0.0	0.3	−0.4
θ	0.0	−0.1	0.0	0.0	0.0	−0.3	0.1	−0.2
η	0.1	0.3	0.0	0.0	0.0	0.2	−0.1	0.2
κ	−0.1	−0.2	0.0	0.0	0.0	−0.3	0.3	−0.2
ρ	−0.1	−0.2	0.0	0.0	0.0	−0.6	0.2	−0.3
$\tilde{\varphi}$	0.1	0.3	0.0	0.0	0.0	0.1	0.5	0.2
β_2	0.4	−0.2	0.2	0.3	−0.6	−1.5	1.1	−0.1
Δ	0.0	−0.1	0.0	0.0	0.0	−0.1	0.0	−0.2
β_4	−0.7	−1.9	0.1	0.6	−0.1	−1.3	−5.0	−3.7
N_{1963}	0.0	0.1	0.0	0.0	0.2	0.3	−0.2	0.0
G_{1963}	0.1	0.0	0.0	0.0	0.0	−0.1	0.6	0.0

Table 6: The effect of parameters on the four empirical paths (numbers refer to elasticities of empirical value wrt. parameter)

Parameters	σ	ϵ	β_1	γ	$\tilde{\kappa}$	θ	η	κ	ρ	$\tilde{\varphi}$	β_2	Δ	β_4	N_{1963}	G_{1963}
Curvature	82.0	2.7	219.3	3.7	1.7	4.2	3.0	47.1	4.8	4.0	12.8	2.4	2142.1	2.3	2.0

Table 7: The "curvature" of deviating from the optimal parameter

We compute the expression on the left. The results are in Table 7 for a 5% shock on the parameter of interest. It shows that the parameters that govern the production function: $(\sigma, \beta_1, \beta_2, \beta_4)$ are the hardest to vary and consequently the ones most precisely identified. The exception is ϵ, the elasticity between low-skill labor and machines, which as Proposition 2 makes clear, does not govern the asymptotic growth of income inequality. $\rho, \theta, \eta, \gamma$ all govern the growth rate of the economy and are weakly identified individually. Increases in φ can to a certain extend be accommodated by changes to N_{1963} and consequently neither is very well-identified. The depreciation of capital Δ is also not well identified because it mostly affects the growth rate of the capital stock which also depends on ρ and θ (equation 188). Given that this parameter is the one estimated outside a common range this is a reassuring finding.

Table 8: Parameters from quantitative exercise for a constant G model

Parameter	σ	ϵ	β_1	γ	θ	ρ	β_2	Δ	β_4	$\tilde{\varphi}$	N_{1963}	G
Value	10.6	39.1	0.51	1.03	1	0.012	0.17	0.007	0.84	1.51	0.16	0.90

7.11.4 Details on the constant G calibration

Table 8 shows the calibrated parameters in the alternative model described in section 4.2 where G is an exogenous constant. To match the data as well as possible, this constant G model requires very high elasticity of substitution across intermediates σ and between low-skill workers and machines in automated firms ϵ. The share of automated products is estimated at $G = 0.9$. Figure 24 reports the two moments not mentioned in section 4.2. This model captures well the trend in labor productivity, but performs worse for the equipment stock to GDP ratio than the baseline model.

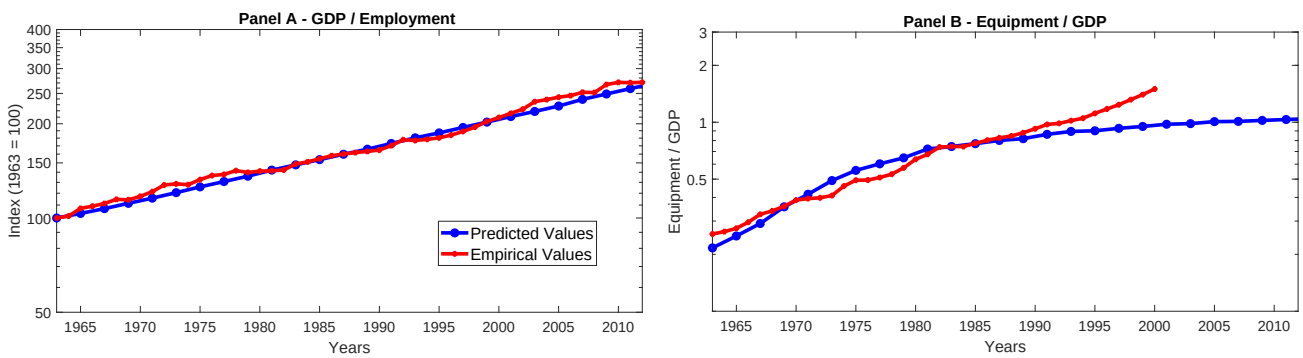

Figure 24: Predicted and empirical time paths for a model with constant G.

7.11.5 Labor-augmenting technical change

In Section 4.2, we showed that conditional on our production function, the data require a path for G_t similar to that generated by our endogenous growth model. To assess how robust that result is, we add high-skill labor augmenting technical change to our model. That is, we replace (22) with:

$$y(i) = [l(i)^{\frac{\epsilon-1}{\epsilon}} + \alpha(i)(\tilde{\varphi} A_{Ht}^{\beta_4} h_e(i)^{\beta_4} k_e(i)^{1-\beta_4})^{\frac{\epsilon-1}{\epsilon}}]^{\frac{\epsilon\beta_1}{\epsilon-1}} A_{Ht}^{\beta_2} h_s(i)^{\beta_2} k_s(i)^{\beta_3},$$

where A_{Ht} is high-skill labor augmenting technical change with g_{A_H} a constant. We then look for the exogenous paths for N_t and G_t together with an initial value A_{H1963} and a growth rate g_{A_H} which best match the data (still assuming the baseline parameters of

Table 1). We find that $A_{H1963} = 1.025$ and a very small $g_{A_H} = 0.03\%$. The resulting path for G_t is still very similar to the one generated by the endogenous growth model as illustrated by Figure 25. Note that allowing instead for exogenous low-skill labor augmenting technical change or machine augmenting technical change (an exogenous trend on $\tilde{\varphi}$) is isomorphic to this exercise for the optimal G_t path (each exercise would deliver a different optimal N_t path).

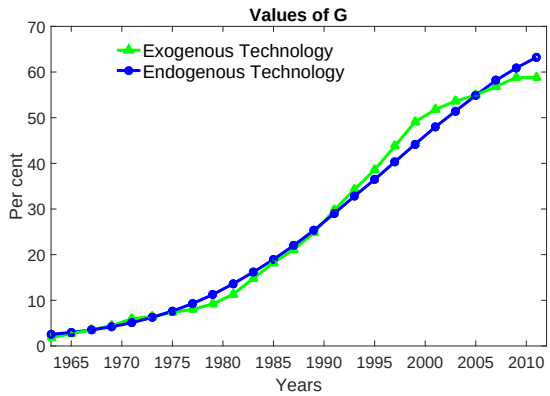

Figure 25: Path $G_t$in the endogenous growth model and in the modified exogenous growth model with high-skill augmenting technical change.

7.11.6 Automation taxes

Among the many policy proposals to address rising income inequality, is a tax on the use of automation technology or a "robot tax". Here, we analyze two distinct taxes: on the use of machines—in the form of a tax on the rental rate of equipment— and on the innovation of new machines—in the form of taxing high-skill workers in automation innovation—see Appendix 7.10 for details. In either cases we consider the permanent unexpected introduction of a 20% tax in the first non-calibrated year, 2013.

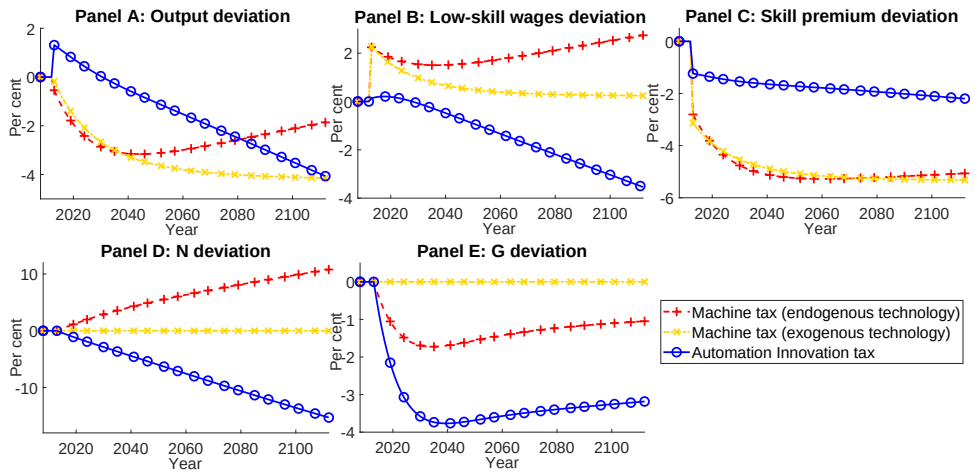

Figure 26: Effects of a machine tax and an automation innovation tax relative to baseline.

First, consider a tax on the use of machines. To clarify the role of endogenous technology we also simulate the economy holding technology, N_t and G_t and therefore H_t^P at the baseline level. Figure 26 reports the results. The immediate effect is to discourage the use of machines and consequently low-skill wages rise by 2% on impact (Panel B) with a corresponding lower skill premium (Panel C) (In Appendix 7.10.5 we show that low-skill wages will increase on impact for any parameter values.[44]) The endogeneity of technology amplifies the effect of the tax over time (in panel B, the gap between the endogenous and the exogenous cases widens). This results from two effects. First, the tax discourages automation innovation leading to a lower G (Panel E). Second, since high-skill workers and machines are complements, the tax reallocates high-skill workers away from production and toward horizontal innovation, increasing N (Panel D). Consequently, the positive effect on low-skill wages is eventually larger than the initial 2%. Output initially decreases on impact in a similar fashion whether technology is endogenous or not, but it recovers and eventually (beyond the horizon of the figure) increases in the endogenous case (Panel A) due to the increase in N_t.[45]

A tax on automation innovation has very different implications: First, high-skill workers move from innovation in automation to production which, on impact, boosts output and marginally low-skill wages. As the share of automated products G_t decreases,

[44]By comparison in KORV, the effect of such a tax depends on parameters.

[45]Asymptotically, a machine tax has no effect on G or on the growth rate of N: as using low-skill workers instead of machines becomes prohibitively expensive, the allocation of high-skill workers remains undistorted by the presence of a finite tax. As a result, in the long-run, G_t reaches the same steady-state but N_t is at a permanently higher level because for a long time the tax has created excess horizontal innovation. See Proposition 10 in Appendix 7.10.

low-skill wages further modestly increase. However, discouraging automation innovation also discourages horizontal innovation since the not-yet automated firms are the ones bearing the burden of the tax. This eventually reduces low-skill wages. The intuition is similar to that of Proposition 7 since a tax on automation innovation has similar effects to reducing the effectiveness of the automation technology. Quantitatively, the effect remains modest since it takes 30 years for the number of products to decrease by 5% (which correspond to a decrease of 0.17 p.p. in annual growth rate). The skill-premium is also reduced as the economy grows at a slower rate.

This exercise highlights the importance of endogenous technology: Though both forms of "robot" taxes increase low-skill wages on impact, the long-run effects depend crucially on whether the tax is designed to encourage or discourage overall innovation. Of course, this exercise is only a first pass and analyzing the welfare consequences of these policies or others, say minimum wage legislation, is of interest for future research.

7.12 Comparison with KORV

We show formally the claims made in Section 4.1 that KORV cannot replicate a decline in the labor share without other counterfactual predictions and does not feature labor-saving innovation. Using their notation, their production function is given by:

$$F = Ak_s^{\alpha}\left(\mu u^{\sigma} + (1-\mu)\left(\lambda k_e^{\rho} + (1-\lambda)\, s^{\rho}\right)^{\sigma/\rho}\right)^{\frac{1-\alpha}{\sigma}}, \tag{189}$$

where k_s is structure, u is low-skill labor, s is high-skill labor and k_e is equipment. The key features are that $\sigma > \rho$ and k_e increases faster than GDP.

In their estimation, k_e and h are strict complements ($\rho < 0$), so as k_e keeps increasing because of investment-specific technological change, its factor share must eventually go to 0; meaning that the long-run prediction of their model is an *increase* in the labor share. Even though, their estimation rejects $\rho \geq 0$, it is worth checking what happens in that case. If equipment and high-skill workers are substitutes ($\rho > 0$, which is the calibrated parameter in Eden and Gaggl, 2018), the economy experiences explosive growth (which seems counterfactual) since in the long-run it becomes an AK model where K/Y grows from investment specific technical change. If $\rho = 0$, then their production function looks like the one of our automated products, and indeed the capital share must eventually

increase. But, then, the growth rate of the skill premium is given by:[46]

$$g_{\pi t} = (1-\sigma)(g_u - g_s) + \sigma\lambda(g_{k_{et}} - g_s).$$

Consequently, if investment specific technological change accelerates (that is there are relatively more and more innovations of that type such that $g_{k_{et}}$ grows), then the skill premium must grow faster (this is also the case for $\sigma > \rho > 0$). This parameterization will now have problems with the first puzzle that we solve: namely a slow down in the growth rate in the skill premium at a time where technical change is the most directed toward "automation" / investment specific technical change.[47]

We now show that investment-specific technical change is not low-skill labor saving in KORV. To do so, we solve for the low-skill wage in their model and consider an increase in investment specific technical change q_t. q_t is the extra TFP parameter in the production of equipment investment compared to the consumption good (so that $1/q_t$ is the price of the investment good for equipment). We look here at the effect of a one time permanent increase in q_t, keeping the expected price change $E_t\left(\frac{q_t}{q_{t+1}}\right)$ fixed and assuming a fixed rental rate on structures (or equivalently a fixed interest rate) $R_{st} = r_t + \delta_s$. Note that we need to make such assumptions because KORV do not specify a supply function for capital (since the capital stock is simply taken from the data). This assumption corresponds to a perfectly elastic capital stock which is how we evaluate the one time effect of a change in G_t in Proposition 1, in the discussion of the effect of automation on wages in Section 4 and in Proposition 8 in Appendix 7.10. Taking first order condition in (189), we get the rental rate on structures:

$$R_{st} = \alpha k_{st}^{\alpha-1}\left(\mu u_t^{\sigma} + (1-\mu)\left(\lambda k_{et}^{\rho} + (1-\lambda)\, s_t^{\rho}\right)^{\frac{\sigma}{\rho}}\right)^{\frac{1-\alpha}{\sigma}}. \tag{191}$$

[46]Here we use their equation (4) $g_{\pi t} = (1-\sigma)(g_u - g_s) + (\sigma-\rho)\lambda\left(\frac{k_{et}}{s_t}\right)^{\rho}(g_{k_{et}} - g_s)$, leaving out the labor augmenting terms, which are not included in their preferred specification and do not reflect capital-skill complementarity. This does not affect the present point.

[47]KORV briefly discuss a production function where the nests are inverted so that:

$$F = Ak_s^{\alpha}\left(\mu s^{\sigma} + (1-\mu)\left(\lambda k_e^{\rho} + (1-\lambda)\, u^{\rho}\right)^{\sigma/\rho}\right)^{\frac{1-\alpha}{\sigma}} \tag{190}$$

with $\sigma < \rho$. This specification does not match their data but is similar to our specification within automated firm (but not for the aggregate economy). Here again the same issues arise: if $\sigma < 0$, then the long-run capital share declines. If $\sigma > 0$, growth is explosive. If $\sigma = 0$ and $\rho > 0$, the capital share increases in the long-run but the skill premium cannot grow less fast when technical change is the most directed toward investment.

KORV assume that the returns on both capital stocks must be equal, that is:

$$1-\delta_s+R_{st}=E_t\left(\frac{q_t}{q_{t+1}}\right)(1-\delta_e)+q_tR_{et}, \tag{192}$$

which implies that the rental rate on equipment obeys:

$$R_{et}=\frac{1}{q_t}\left(1-\delta_s+R_{st}-E_t\left(\frac{q_t}{q_{t+1}}\right)(1-\delta_e)\right).$$

Therefore R_{et} decreases with q_{et}. Taking first order condition in (189) with respect to k_{et}, and using (191), we get:

$$R_{et} \;=\; (1-\mu)\,\lambda\left(\lambda+(1-\lambda)\frac{s_t^\rho}{k_{et}^\rho}\right)^{\frac{1-\rho}{\rho}}(1-\alpha)\left(\frac{\alpha}{R_{st}}\right)^{\frac{\alpha}{1-\alpha}}\left(\mu\frac{u_t^\sigma}{\left(\lambda k_{et}^\rho+(1-\lambda)\,s_t^\rho\right)^{\frac{\sigma}{\rho}}}+(1-\mu)\right)^{\frac{1-\sigma}{\sigma}},$$

which shows that (as expected) k_{et} decreases in R_{et} so that k_{et} increases if q_t increases. Finally, the first order condition with respect to unskilled labor is given by

$$w_{Lt}=(1-\alpha)\,k_{st}^\alpha\mu u_t^{\sigma-1}\left(\mu u_t^\sigma+(1-\mu)\left(\lambda k_{et}^\rho+(1-\lambda)\,s_t^\rho\right)^{\frac{\sigma}{\rho}}\right)^{\frac{1-\alpha}{\sigma}-1}. \tag{193}$$

Combining this with (191) gives

$$w_{Lt} \;=\; (1-\alpha)\left(\frac{\alpha}{R_{st}}\right)^{\frac{\alpha}{1-\alpha}}\mu u_t^{\sigma-1}\left(\mu u_t^\sigma+(1-\mu)\left(\lambda k_{et}^\rho+(1-\lambda)\,s_t^\rho\right)^{\frac{\sigma}{\rho}}\right)^{\frac{1-\sigma}{\sigma}}. \tag{194}$$

Therefore an increase in q_t leads to an increase in k_{et} and consequently low-skill wages w_{Lt}: investment specific technical change is not labor saving in KORV's main specification.[48] This is true regardless of the parameters σ and ρ (and therefore also applies to Eden and Gaggl, 2018).

[48] It can be labor-saving with the alternative nesting given in (190).