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

David Hémous and Morten Olsen *

August 2019 (first draft: September 2013)

Abstract

We construct an endogenous growth model with automation and horizontal innovation in an economy with low- and high-skill workers. Automation enables the replacement of low-skill workers with machines, increasing the skill premium and decreasing the labor share. Horizontal innovation increases both wages. The share of automation innovations endogenously increases over time through an increase in low-skill wages. We calibrate the model to the US economy and show that it replicates the paths of the skill premium, the labor share and labor productivity. Further, taxing automation innovation increases low-skill wages in the short run but reduces them in the long run.


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 Daron Acemoglu, Philippe Aghion, Ufuk Akcigit, Pol Antrás, Tobias Broer, Steve Cica, Per Krusell, Brent Neiman, Jennifer Page, Ofer Setty, Andrei Shleifer, Che-Lin Su, Fabrizio Zilibotti and Joachim Voth among others for helpful comments and suggestions. We also thank seminar and conference participants at IIES, University of Copenhagen, Warwick, UCSD, UCLA Anderson, USC Marshall, Barcelona GSE Summer Forum, the 6th Joint Macro Workshop at Banque de France, Chicago Harris, the 2014 SED meeting, the NBER Summer Institute, the 2014 EEA meeting, Ecole Polytechnique, the University of Zurich, NUS, London School of Economics, the 2015 World Congress of the Econometric Society, ECARES, Columbia University, EIEF, 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 last 50 years, the United States has seen dramatic changes in the income distribution. The skill premium has increased by 33% between 1963 and 2012 and the labor share has declined by 7 p.p. since the 1970s (Figure 1A and B). Meanwhile, several automation technologies (numerically controlled machine tools, automatic conveyor systems, industrial robots,...) were 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 1C plots the share of automation patents in the US according to Mann and Püttmann, 2018).

In this paper, we develop a model where the rise in the skill premium and the decline in the labor share result from an endogenous increase in the share of automation innovations. Our model combines two forms of technological change: Horizontal innovation, modeled as in Romer (1990), increases demand for both high-skill and low-skill workers. Automation innovation takes place in existing product lines and enables the replacement of low-skill workers in production with machines. Consequently, our model features a task framework where machines can substitute for workers as Autor, Levy and Murnane (2003) and directed technical change as Acemoglu (1998) since innovation endogenously occurs in two different technologies. Furthermore, it embodies capital-skill complementarity as Krusell, Ohanian, Ríos-Rull and Violante (2002, henceforth KORV). Whereas 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 current trends in inequality and factor shares, we argue instead that those should be seen as the natural progression of an economy: in our model, the share of automation innovations increases as the economy grows.

Moreover, the interplay between automation and horizontal innovation allows us to account for two salient puzzles in the literature: the stagnation of labor productivity growth despite the rise in automation innovations and the decline in the growth rate of the skill premium since the mid 1990s without an apparent decline in skill-biased technological change. Furthermore, we show that an economy with endogenous technical change responds differently to policy interventions compared to an economy with exogenous technological progress.

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, an increase in
the level of automation in the economy has an ambiguous effect on low-skill wages, but
we derive conditions under which automation is low-skill labor-saving. Moreover, an
increase in automation always increases the skill premium and reduces the labor share,
in line with the recent evolution of the income distribution. Then, we study general
exogenous processes of horizontal and automation innovations. Over time, technological
progress involves the creation of new (non-automated) tasks for low-skill workers, who
are then subsequently replaced by machines. Nevertheless, this is not enough to ensure
that low-skill and high-skill wages grow at the same rate: the asymptotic growth rate of
low-skill wages must be positive but strictly lower than that of high-skill wages.

Second, we endogenize innovation, which allows us to rationalize the observed in-
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 and therefore the share of automation
innovation increase. 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.

In a third step, 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 the 1960s. For this exercise, we identify skill groups with education groups, such

---

1 For instance, the introduction of the telephone led to the creation of new jobs. In 1970 there were
421 000 switchboard operators in the United States. This occupation has largely been automated today.
that high-skill workers correspond to college-educated workers. Our model reproduces the trends in the data quantitatively. In particular, labor productivity growth stagnates because horizontal innovation declines and the growth rate of the skill premium declines in the 1990s and 2000s in spite of more innovation being directed toward automation.

Intuitively, this comes from looking at automation as a stock: with a higher share of automated products, there must be more automation innovations to compensate for its depreciation through horizontal innovation. Our model also performs well in out-of-sample predictions: a calibration only over the years 1963-1993 gives very similar results for the years 1994-2012. Finally, we use the calibrated model for policy experiments. A tax on the use of machines has a positive effect on low-skill wages which is amplified by the endogeneity of technology. In contrast, a tax on automation innovation initially increases low-skill wages but ends up having a negative impact after a few years.

Our modeling of automation as high-skill-biased is motivated by a large empirical 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 Acemoglu and Restrepo, 2017b) decrease the relative demand for low-skill workers.

A large macro literature has argued that skill-biased technical change 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, and Beaudry and Green, 2005). While such theories explain transitory increases in inequality and ignore the dynamics of factor shares, our model features widening inequality and a decline in the labor share. 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

---

2Therefore, our model addresses the Card and DiNardo (2002) critique that the slow-down in the skill premium is inconsistent with a the skill-biased technological change hypothesis.

3Autor, Katz, and Kearney (2006, 2008) and Autor and Dorn (2013) relate wage and job polarization with the computer-driven automation of routine tasks often performed by middle-skill workers. For our purpose, we will not distinguish between low- and middle-skill workers since both have often performed tasks which have later on been automated (a previous version of this paper, Héamous and Olsen (2016) did so). See also Feng and Graetz (2016).

4Beaudry, Green and Sand (2016) model the IT revolution as an exogenous increase in the demand for organizational capital built by cognitive labor. Once the capital stock reaches steady-state, the demand for cognitive tasks decreases. We obtain a similar pattern in the growth rate of the skill premium.
variation in the skill premium. Our model also features capital-skill complementarity but differs in several dimensions. In particular, our model features 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, building on Katz and Murphy (1992), considers technology to be low-skill or high-skill labor augmenting. Goldin and Katz (2008) find that technical change in the US was skill-biased throughout the 20th century, and Katz and Margo (2014) that the relative demand for white-collar workers has been increasing since 1820. Further, the directed technical change literature (Acemoglu, 1998, 2002, 2007) endogenizes the bias of technology. Such models have no role for labor-replacing technology and therefore cannot generate changes in the labor share (see Acemoglu and Autor, 2011).

The idea that high wages might incentivize technological progress in the form of labor-saving innovations dates back to Habakkuk (1962). In Zeira (1998), exogenous increases in TFP raise wages and encourage the adoption of a capital-intensive technology, which further raises wages. 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. Our solution to obtain a more realistic path for wages is to introduce a second type of technological progress, namely the creation of new products 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. Yet, 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 their focus is on the self-correcting elements of the economy after a technological shock. In contrast, our model does not feature a balanced growth path and we focus on accounting for secular trends. Subsequent papers combining automation and horizontal innovations include Martinez (2018) and Rahman (2017).

A few recent papers provide empirical evidence for the role of wages on technology adoption: Lewis (2011), Hornbeck and Naidu (2014) and Clemens, Lewis and Postel (2018), who study the consequences of migration shocks, Manuelli and Seshadri (2014) who study the adoption of tractors and Acemoglu and Restrepo (2018b) who relate robot

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

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 and conducts policy exercises. Section 5 concludes. The Main Appendix presents the main extensions of our model and discusses the identification of our parameters. The Secondary Appendix presents proofs, other extensions and details the calibration exercise.

2 A Baseline Model with Exogenous Innovation

This section presents a baseline 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. Section 2.5 discusses some of our assumptions.

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 $\rho$ 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 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:

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

(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:

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

(2)

where $\alpha(i) \in \{0, 1\}$ is an indicator function for whether or not the monopolist has access to an automation technology which allows for the use of machines. If the product is not automated ($\alpha(i) = 0$), production takes place using a Cobb-Douglas production function with only low-skill and high-skill labor and a low-skill factor share of $\beta$. If it is automated ($\alpha(i) = 1$) machines can be used in the production process. We allow for perfect substitutability, in which case $\epsilon = \infty$ and the production function is $y(i) = [l(i) + \alpha(i)\bar{\varphi}x(i)]^{\beta} h(i)^{1-\beta}$. The parameter $\bar{\varphi}$ 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.

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 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, secondary innovations in a growth model were introduced by Aghion and Howitt (1996) who study the interplay between applied and fundamental research.
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\textsuperscript{7}

For now, machines are an intermediate input. Once invented, machines of type $i$ are produced competitively one for one with the final good, such that the price of an existing machine for an automated firm 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 technological levels $N$ (the number of products) and $G$ (the share of automated products) 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).

First, note that all automated firms are symmetric and therefore behave in the same way. Similarly all non-automated firms are symmetric. This gives aggregate output of:

$$Y = N^{\frac{1}{1-\sigma}} \times \left( (1-G)^{\frac{1}{\sigma}} \left( \left( \frac{L^A}{T_1} \right)^{1-\beta} \right)^{\frac{1}{\sigma}} + G^{\frac{1}{\sigma}} \left( \left( \left[ \frac{L^A}{T_2} + \frac{\tilde{\phi} X}{T_2} \right]^{-\frac{1}{1-\epsilon}} \right)^{\frac{1}{\sigma}} \right)^{\frac{1}{\sigma}} \right)^{\frac{1}{\sigma}} \right),$$

where $L^A$ (respectively $L^NA$) is the total mass of low-skill workers in automated (respectively non-automated) firms, $H^{P,A}$ (respectively $H^{P,NA}$) is the total mass of high-skill workers hired in production in automated (respectively non-automated) firms and $X = \int_0^N x(i) \, di$ is total use of machines. The aggregate production function takes the form of a nested CES between two sub-production functions. The first term $T_1$ captures the classic case where production takes place with constant shares between factors (low-skill and high-skill labor). 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 $\sigma > 1$). $N^{\frac{1}{1-\sigma}}$ is a TFP parameter. The aggregate production function \textsuperscript{[3]} differs from the typical aggregate CES production function with factor-augmenting technical change because we microfound it with a model of tasks’ automation. Moreover,

\textsuperscript{7}We include IT innovations in our interpretation because our model does not distinguish between low-skill and middle-skill workers.
changes in the share parameter $G$ will have different effects on wages compared to the typical changes in factor-augmenting technologies.

The unit cost of product $i$ is given by:

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

where $\varphi \equiv \tilde{\varphi}^\epsilon$, $w_L$ denotes low-skill wages and $w_H$ high-skill wages. When $\epsilon < \infty$, $c(\cdot)$ is strictly increasing in both $w_L$ and $w_H$ and $c(w_L, w_H, 1) < c(w_L, w_H, 0)$ for all $w_L, w_H > 0$ (automation reduces costs). 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 (4) delivers the demand for low-skill labor of a single firm.

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

which decreases in $w_L$ and $w_H$. 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 and ensures $l(w_L, w_H, 1) < l(w_L, w_H, 0)$ for all $w_L, w_H > 0$.

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

$$x(w_L, w_H)/l(w_L, w_H, 1) = \varphi w_L^\epsilon, \quad (6)$$

which increases in $w_L$ as the wage is also the price of low-skill labor relative to machines.

The iso-elastic demand [1], coupled with constant mark-up $\sigma/(\sigma - 1)$, implies that revenues are given by $R(w_L, w_H, \alpha(i)) = ((\sigma - 1)/\sigma)^{\sigma-1} c(w_L, w_H, \alpha(i))^{1-\sigma} Y$ and profits are a fixed share of revenue: $\pi(w_L, w_H, \alpha(i)) = R(w_L, w_H, \alpha(i))/\sigma$. We define $\mu \equiv \beta(\sigma - 1)/\epsilon < 1$ (by our assumption on $\epsilon$). Using (4), the relative revenues (and

---

8For 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.
profits) of non-automated and automated firms are given by:

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

(7)

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

Since firms’ profits are a constant share of firms’ revenues, aggregate profits are a constant share \( 1/\sigma \) of output \( Y \). Similarly, the share of firms’ revenues accruing to high-skill labor in production is the same for all firms and given by \( \nu_h = (1 - \beta)(\sigma - 1)/\sigma \). Therefore payment to high-skill labor in production is a constant share of output:

\[
w_H H = (1 - \beta)\frac{\sigma - 1}{\sigma} N [GR(w_L, w_H, 1) + (1 - G)R(w_L, w_H, 0)] = (1 - \beta)\frac{\sigma - 1}{\sigma} Y. \tag{8}
\]

Using factor demand functions, the share of revenues accruing to low-skill labor is given by \( \nu_l(w_L, w_H, \alpha(i)) = \frac{\sigma - 1}{\sigma} \beta (1 + \varphi w_L^{\epsilon - 1} \alpha(i))^{-1} \), and decreases with automation. Using labor market clearing \( \int_0^N l(i)di = L \), we obtain total wages of low-skill workers as:

\[
w_L L = N [GR(w_L, w_H, 1)\nu_l(w_L, w_H, 1) + (1 - G)R(w_L, w_H, 0)\nu_l(w_L, w_H, 0)]. \tag{9}
\]

Equations (7), (8) and (9) give the high-skill to low-skill labor share in production as:

\[
\frac{w_H H^P}{w_L L} = \frac{1 - \beta}{\beta} G + \frac{G + (1 - G)(1 + \varphi w_L^{\epsilon - 1})^{-\mu}}{(1 + \varphi w_L^{\epsilon - 1})^{-1} + (1 - G)(1 + \varphi w_L^{\epsilon - 1})^{-\mu}}. \tag{10}
\]

This expression gives the relative demand curve for high-skill and low-skill labor. We represent this relationship for given technology levels and factor supply ratio \( L/H^P \) by a curve in the \((w_L, w_H)\) space in Figure 2. For \( G = 0 \), the 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 (10) increases in \( w_L \), so that the relative demand curve is non-homothetic and rotates counter-clockwise as \( w_L \) grows. Intuitively, higher low-skill wages increase the ratio of high-skill to low-skill labor share in production because they induce both more substitution toward machines in automated firms (as reflected by the term \((1 + \varphi w_L^{\epsilon - 1})^{-1}\) in (10)) and improve the cost-advantage of automated firms (term \((1 + \varphi w_L^{\epsilon - 1})^{-\mu}\) ). Therefore, as long as \( G \) tends toward a positive constant, low-skill and

\[9\]

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

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

$$\frac{\sigma}{\sigma - 1} \frac{N^{\frac{1}{\sigma}}}{\beta^\beta (1-\beta)^{1-\beta}} \left( G \left( \varphi + w_L^{1-\epsilon} \right)^\mu + (1-G)w_L^{\beta(1-\sigma)} \right) \frac{1}{1-\sigma} w_H^{1-\beta} = 1. \quad (11)$$

This relationship defines the *unit isocost curve* 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. Together (10) and (11) determine real wages uniquely as a function of $N, G$ and $H_P$.

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

$$Y = C + X \quad (12)$$

where $C = C_L + C_H$ is total consumption. GDP includes the payment to labor and aggregate profits so GDP and the total labor share are given by

$$GDP \equiv \frac{1}{\sigma} Y + w_L L + w_H H, \quad LS = 1 - \frac{1}{1 + (\sigma - 1) (1-\beta) \left( \frac{w_L L}{w_H H_P} + \frac{H}{H_P} \right)}, \quad (13)$$

where the second equality uses (8). Therefore, the labor share decreases in the skill premium for a given mass of high-skill workers in production $H_P$.  

10
2.3 Technological change and wages

We analyze the consequences of technological 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 types of 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 productivity 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 productivity effect. Formally, one can show (proof in Appendix 7.1)\(^{10} \)

**Proposition 1.** Consider the equilibrium \((w_L, w_H)\) determined by equations \((10)\) and \((11)\). 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 four cases: i) The

\(^{10}\)In the perfect substitute case, \( \epsilon = \infty \), \( w_H \) increases in \( N \) and weakly increases in \( G \), \( w_H/w_L \) 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.
elasticity of substitution $\sigma$ is large, as newly automated products gain a larger market share. \textit{ii)} The cost share of the low-skill labor-machines aggregate $\beta$ is small as the cost-saving effect of automation is small. \textit{iii)} $G$ is high, since in that case there are few non-automated firms, so that the automation of one more firm hurts low-skill workers more, while most of the aggregate productivity gains are realized for low $G$. \textit{iv)} $N$ is large since higher wages make the substitution effect stronger. It is worth comparing the effect of automation with that of an increase in machines’ productivity $\varphi$ (equivalent to a decline in the price of machine). An increase in $\varphi$ also has an ambiguous effect on low-skill wages resulting from the combination of a substitution effect and a scale effect. Yet, we can show that an increase in machine’s productivity is less likely to be low-skill labor saving than an increase in the share of automated products.\textsuperscript{11} In contrast to our paper, neither a factor-augmenting technical change model with directed technical change as Acemoglu (1998), nor a capital-skill complementarity model like KORV (with capital supplied at a constant price like our machines here) feature labor-saving technical change.

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. In Appendix \textsuperscript{7.1} we show that horizontal innovation is low-skill labor biased (it reduces the skill premium) if and only if $1 - \beta > ((\epsilon - 1) / (\sigma - 1) - \beta) / (1 + \varphi w_L^{-1})$. This condition is met when $w_L$ is high enough (that is for $N$ large enough) in which case the isocost curve does not move much with horizontal innovation; it is also met for any value of $w_L$ if machines and low-skill workers are not too substitute (for $\epsilon \leq \sigma$).\textsuperscript{12}

Therefore, the secular increase in the skill premium in the US economy (Stylized fact 1) is consistent with an economy with a growing number of products and a constant or rising share of automated products. As the labor share is decreasing in the skill premium, its decline (Stylized fact 2) is also consistent with such a pattern of technological

\textsuperscript{11}Formally, we show that $\frac{\partial c}{\partial \varphi} < 0$ implies that $\frac{\partial w}{\partial \varphi} < 0$ but the reverse is not true—see Appendix \textsuperscript{7.1}. 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 (3) and is made by Aghion, Jones and Jones (2017).

\textsuperscript{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.
Furthermore, for some parameters, an increase in the share of automated products could explain a temporary decline in low-skill wages. 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 \( g^a_t \equiv \dot{a}_t/a_t \) denote its growth rate and \( g^a_\infty = \lim_{t \to \infty} g^a_t \) if it exists. In Appendix 7.2.1 we derive:

**Proposition 2.** Consider three processes \([N_t]_{t=0}^\infty\), \([G_t]_{t=0}^\infty\) and \([H_t^P]_{t=0}^\infty\) where \((N_t, G_t, H_t^P) \in (0, \infty) \times [0, 1] \times (0, H)\) for all \( t \). Assume that \( G_t, g^N_t \) and \( H_t^P \) all admit limits \( G_\infty, g^N_\infty \) and \( H^P_\infty \) with \( g^N_\infty > 0 \) and \( H^P_\infty > 0 \).

A) If \( G_\infty \in (0, 1) \), the asymptotic growth of high-skill wages \( w_{Ht} \) and output \( Y_t \) are:

\[
g^w_{H\infty} = g^Y_\infty = g^N_\infty / ((1 - \beta)(\sigma - 1)) ,
\]

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

\[
g^{wL}_\infty = g^Y_\infty / (1 + \beta(\sigma - 1)) .
\]

B). If \( G_\infty = 1 \), the asymptotic growth rates of \( w_{Ht} \) and \( Y_t \) also obey \(14\). If \( G_t \) converges sufficiently fast (such that \( \lim_{t \to \infty} (1 - G_t) N_t^{\psi(1-\mu)-1} \) exists and is finite) then :

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

\[
g^{wL}_\infty = g^Y_\infty / \epsilon .
\]

- 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 \( \lim_{t \to \infty} (1 - G_t) N_t^{\psi} = 0 \)).

C) 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:

\[
g^{wL}_\infty = g^{wH}_\infty = g^Y_\infty = g^N_\infty / (\sigma - 1).
\]

\[\]
This proposition first relates the growth rates of output and high-skill wages to the growth rate of the number of products. Without automation, that is if $G_t$ converges to 0 sufficiently fast, $Y_t$ is proportional to $N_t^{1/(\sigma-1)}$ as in a standard expanding-variety model. Automation introduces machines as an additional reproducible input such that a higher level of productivity leads to a higher supply of machines further increasing output when $G_\infty > 0$. This multiplier effect is increasing in the asymptotic share of machines, $\beta$.

Second, with positive growth in $N_t$, mild assumptions are sufficient for asymptotic positive growth in low-skill wages. $w_{Lt}$ only remains bounded when there is economy-wide perfect substitution, i.e. low-skill workers and machines are perfect substitutes, $\epsilon = \infty$, and all products are automated asymptotically ($G_t$ converges to 1 sufficiently fast). Even then, low-skill wages are bounded below by $\bar{\varphi}^{-1}$, as a lower wage would imply that no firm would use machines. In general, the processes of $N_t$ and $G_t$ depend on the rate at which new products are introduced, the extent to which they are initially automated, and the rate of automation. 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.

With aggregate imperfect substitution (because $G_\infty < 1$ or $\epsilon < \infty$), 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 ($p^*_x$) must grow at a positive rate. Since machines are produced with the same technology as the consumption good, $p^*_x = p^C_t$, where $p^C_t$ is the price of the consumption good (1 with our normalization), the real wage $w_{Lt} = w_{Lt}/p^C_t = (w_{Lt}/p^*_x)(p^*_x/p^C_t)$ must also 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 is easily seen from the aggregate production function (3), which asymptotically is a nested CES with constant share parameters and where technological change in the form of an increase in $N_t$ is not labor-augmenting unless $G_\infty = 0$. The result then follows from Uzawa’s theorem. We get $G_\infty > 0$ as long as the automation intensity is bounded away from 0 (see Appendix 7.2.2).

If $\epsilon < \infty$ and $G_\infty = 1$ (sufficiently fast), low-skill workers 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, $\epsilon$.

In contrast, when $G_\infty \in (0,1)$, the demand for low-skill labor increasingly comes
from the non-automated firms (as automation is labor-saving at the firm level). With
growing wages, the relative market share of non-automated firms decreases in proportion
with \((1 + \varphi w^{-1}_{Lt})^{-\mu} \sim \varphi^{-\mu} w^{-\beta(\sigma-1)}_{Lt}\), while most of the demand for high-skill labor comes
from automated firms. Then, the growth rate of low-skill wages is a fraction of the
growth rate of high-skill wages given by \((15)\). The ratio between high-skill and low-skill
wage growth rates increases with a higher importance of low-skill workers (a higher \(\beta\)) or
a higher substitutability between automated and non-automated products (a higher \(\sigma\))
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, \(\epsilon\) or
of the exact asymptotic share of automated products \(G_\infty\). 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 \((15)\) instead of \((16)\) and \(\epsilon > 1 + \beta(\sigma - 1)\)). In the model of section \([3]\) the economy endogenously ends up in this case.

2.5 Discussion

Proposition \([2]\) establishes general conditions under which low-skill wages asymptotically
grow but slower than high-skill wages. We now discuss the robustness of this result.
First, our assumption that machines are an intermediate input is innocuous: Section \([4]\)
relaxes this assumption and lets machines take the form of capital with no qualitative
change of result. Second, Appendix \([7.10]\) relaxes the assumption of an exogenous stock
of labor and considers a Roy model where workers are heterogeneous in the quantity
of high-skill labor they can supply. Proposition \([2]\) generalizes to this case, although the
relative growth rate of low-skill wages is higher and asymptotically all workers supply
high-skill labor. Third, Appendix \([7.4]\) presents a model where the production technologies
for machines and the consumption good differ allowing for negative growth in \(p^e_t/p^C_c\). In
this case, low-skill wages may decline asymptotically. Fourth, even if some of the tasks
(but not all) performed by high-skill workers are automatable, our results would remain
similar as long as high-skill workers remain essential in production\([14]\).

Appendix \([7.3]\) breaks the assumption of symmetry and assumes that new products
have a higher TFP or a higher labor productivity. We show that low-skill wages can only
grow at the same rate as output if all productivity improvements are labor-augmenting

\[14\] If all labor tasks are automatable, infinite production is possible in finite time once \(N\) is large
enough. Factors such as natural resources or land are then likely to be the scarce factor.
(as in Acemoglu and Restrepo, 2017a). As soon as potential machines are also at least partly more productive in new tasks, high-skill wages grow faster than low-skill wages.\footnote{We have also abstracted from the accumulation of low-skill human capital. “Traditional” human capital would be equivalent to augmenting low-skill labor in (3), and, from Uzawa’s theorem, would be insufficient to guarantee that low-skill wages grow at the same rate as output when $G_{\infty} > 0$ and $N_t$ grows exponentially. Grossman et al. (2017) show how non-traditional human capital can lead to balanced growth when technological change is not purely labor-augmenting.}

3 Endogenous innovation

We now model automation and horizontal innovation as the result of investment. This allows us first to look at the impact of wages on technological change (the reverse of Proposition 1), second to 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, and third to explore the interactions between the two innovation processes. Section 3.1-3.4 characterizes the model and its solution and section 3.5 provides additional results and comparative statics.

3.1 Modeling innovation

If a non-automated firm hires $h_A^t(i)$ high-skill workers to perform automation research, it becomes automated at a Poisson rate $\eta\tilde{G}_t^\kappa (N_t h_A^t(i))^\kappa$. Once a firm is automated it remains so forever. $\eta > 0$ denotes the productivity of the automation technology, $\kappa \in (0, 1)$ measures the concavity of the automation technology, $G_t^\kappa$, $\tilde{\kappa} \in [0, \kappa]$, represents possible knowledge spillovers from the share of automated products, and $N_t$ represents knowledge spillovers from the total number of 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.\footnote{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 $Y_t = \int_0^{N_t(j)} y_t(i,j)^{\frac{1}{\sigma}} didj$. When a firm hires $H_t^A(j)$ high-skill workers in automation each of its non-automated products gets independently automated with a Poisson rate of $\eta G_t^\kappa (H_t^A(j)/(1 - G_t(j)))^\kappa$. 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).} The presence of spillovers in automation technology ($\tilde{\kappa} > 0$) implies a
delayed and faster rise in the share of automated products. We assume that \( \tilde{\kappa} < 1 - \kappa(1 - \beta) \) which ensures that automation always takes off (see Proposition 4).

New products are developed by high-skill workers in a standard manner according to a linear technology with productivity \( \gamma N_t \). 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 (following Proposition 1) is low-skill biased under certain conditions. This is motivated by the idea that when a task is new and unfamiliar, the flexibility and outside experience of workers allow them to solve unforeseen problems. Only as the task becomes routine and potentially codeifiable, a machine (or an algorithm) can perform it (Autor, 2013).

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

\[
\dot{G}_t = \eta G_t \tilde{\kappa} (N_t h_t^A)^\kappa (1 - G_t) - G_t g_t^N.
\]

Therefore, the level of automation in the economy, \( G_t \), can be understood as a “stock” that gets depreciated 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 \( (\eta G_t \tilde{\kappa} (N_t h_t^A)^\kappa (1 - G_t)) \) is required to raise 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 rise the fastest when innovation is the most directed toward automation. It is also one of the main differences between our modeling of automation 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. We define the total mass of high-skill workers working in automation as \( H_t^A \equiv \int_0^{N_t} h_t^A(i)di \).

\( ^{17} \) Whenever \( \tilde{\kappa} > 0 \), 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.

\( ^{18} \) Using that by symmetry the total amount of high-skill workers hired in automation research is...
skill labor market clearing then leads to

\[ H^A_t + H^D_t + H^P_t = H. \]  

(19)

3.2 Innovation allocation

We denote by \( V^A_t \) the value of an automated firm, by \( r_t \) the economy-wide interest rate and by \( \pi^A_t \equiv \pi(w_{Lt}, w_{Ht}, 1) \) the profits at time \( t \) of an automated firm. The asset pricing equation for an automated firm is given by

\[ r_t V^A_t = \pi^A_t + \dot{V}^A_t. \]  

(20)

This equation states that the required return on holding an automated firm, \( V^A_t \), 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 \( V^N_t \) the value of a non-automated firm and letting \( \pi^N_t \equiv \pi(w_{Lt}, w_{Ht}, 0) \), we get the asset pricing equation:

\[ r_t V^N_t = \pi^N_t + \eta \tilde{G}^\kappa_t (N_t h^A_t) - w_{Ht} h^A_t + \dot{V}^N_t, \]  

(21)

where \( h^A_t \) is the mass of high-skill workers in automation research hired by a single non-automated firm. This equation is similar to equation \( (20) \), but profits are augmented by the instantaneous expected gain from innovation \( \eta \tilde{G}^\kappa_t (N_t h^A_t) (V^A_t - V^N_t) \) net of expenditure on automation research, \( w_{Ht} h^A_t \). This gives the first order condition:

\[ \kappa \eta \tilde{G}^\kappa_t N_t (h^A_t)^{\kappa-1} (V^A_t - V^N_t) = w_{Ht}. \]  

(22)

\( h^A_t \) increases with the difference in value between automated and non-automated firms, and thereby current and future low-skill wages—all else equal\(^{19}\).

Free-entry in horizontal innovation guarantees that the value of creating a new firm

\[ H^A_t = (1 - G_t) N_t h^A_t, \]

we can rewrite \( 18 \) as \( \dot{G}_t = \eta \tilde{G}^\kappa_t (H^A_t)^\kappa (1 - G_t)^{1-\kappa} - G_t g^N_t \).

\(^{19}\) 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).
cannot be greater than its opportunity cost:

$$\gamma N_t V_t^N \leq w_{Ht},$$  \hspace{1cm} (23)$$

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

The low-skill and high-skill representative households’ problems are standard and lead to Euler equations which in combination give$^{20}$

$$\dot{C}_t / C_t = (r_t - \rho) / \theta,$$  \hspace{1cm} (24)$$

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 Equilibrium characterization

In Appendix 6.1, we show 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 ((38), (39), (46), (47), (52) and (54)-(56) with normalized variables defined in (33)-(37) in Appendix 6.1). We establish:

**Proposition 3.** Assume that

$$\kappa^{-\kappa} \left( \gamma(1 - \kappa)/\rho \right)^{\kappa-1} \rho/\eta + \rho/\gamma < \psi H.$$  \hspace{1cm} (25)$$

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 $g_\infty^N > 0$ and a positive constant mass of high-skill workers in automation research $H_\infty^A > 0$. The growth rates of output and wages are given by Proposition 2.A as (14) and (15). The growth rates of profits of automated and non-automated firms are given by

$$g_\infty^A = g_\infty^Y - g_\infty^N \text{ and } g_\infty^{\pi_A} = g_\infty^A - \beta (\sigma - 1) g_\infty^{w_L}.$$  \hspace{1cm} (26)$$

$^{20}$Consumption growth is the same for both households even though, as we will see, 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.
and the growth rates of the value of firms (automated or not) are given by

\[ g^V_A = g^V_N = g^{\pi A} = g^Y - g^N. \] (27)

Proposition 3 establishes the existence of an asymptotic steady-state—"asymptotic" because the system of differential equations only admits a fixed point for \( N_t = \infty \). In addition, the assumption that \( \theta \geq 1 \) ensures that the transversality condition always holds. For the rest of the paper we restrict attention to parameters such that there exists a unique saddle-path stable steady state. Then, for an initial pair \((N_0, G_0) \in (0, \infty) \times [0, 1]\) sufficiently close to the asymptotic steady state, the model features a unique equilibrium converging towards it.

Proposition 3 further shows that in the asymptotic steady-state the share of automated products is between 0 and 1 such that Proposition 2A applies: high-skill wages grow at the same rate as output and low-skill wages grow at a positive but lower rate. Further, as high-skill wages grow at a positive rate in the long-run, the profits of non-automated firms grow less fast than those of automated firms. Since total profits are proportional to output, the profits of each automated firm grow at the rate of output minus the growth rate of the number of products. The value of an automated firm corresponds to its discounted flow of profits and therefore grows at the same rate as its profits. Asymptotically, the profits of non-automated firms are negligible relative to those of automated firms, so that the value of a non-automated firm entirely depends on the fact that it becomes automated at a constant Poisson rate, and it grows at the same rate as the value of an automated firm. In other words, it is the prospect of future automation which guarantees the entry of new products—this plays an important role in the interaction between the two innovation technologies.

\[ ^{21} \text{Technically, there is a steady state for the transformed system in Appendix 6.1, where the number of product } N_t \text{ is replaced by an inversely related variable } n_t, n_t = 0 \text{ in steady state.} \]

\[ ^{22} \text{To understand equation (25), let the efficiency of the automation technology } \eta \text{ be arbitrarily large such that the model approaches a Romer model with only automated firms. Then equation (25) becomes } \rho/\gamma < \psi H, \text{ which mirrors the condition for growth in a Romer model with linear innovation technology. With a smaller } \eta \text{ the present value of a new product is reduced and the condition is more stringent.} \]

\[ ^{23} \text{Multiple asymptotic steady states with } G^* > 0 \text{ are technically possible but are not likely for reasonable parameter values (see Appendix 7.5.1). In addition, with two state variables } (n_t \text{ and } G_t) \text{ saddle path stability requires exactly two eigenvalues with positive real parts. In our numerical investigation, for all parameter combinations which satisfy the previous restrictions, this condition was always met.} \]

\[ ^{24} \text{The economy would not admit such an asymptotic steady-state if automation was entirely undertaken by entrants replacing the incumbents. If that were the case, the value of a firm created through horizontal innovation would only correspond to the discounted flow of profits of a non-automated firm, which grows slower than the cost of horizontal innovation (high-skill wages normalized by } N_t). \text{ As} \]
3.4 Innovation incentives along the transitional path

We now describe the equilibrium and explain how the economy reaches the asymptotic steady-state even when starting far from it. In Appendices 7.5.2-7.5.4, we establish:

**Proposition 4.**

A. There exists an $N_0$ sufficiently low such that there is an interval $(0, \hat{t})$ during which the automation rate $\eta G_t (h_t^A)^\kappa$ is small and the economy behaves arbitrarily close to that of a Romer model where automation is impossible.

B. If $g_t^N$ admits a positive limit, $G_t$ cannot converge toward 0. If $G_t$ admits a limit, then the economy converges to an asymptotic steady-state as described in Proposition 3.

This proposition establishes that if $G_t$ admits a limit and $g_t^N$ admits a positive limit then the economy must converge toward the asymptotic steady-state regardless of the initial values ($N_0, G_0$). Moreover, the proposition shows that 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. Appendix 6.2 undertakes a numerical simulation to illustrate this section.

Following (22), 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 $\eta G_t (H_t^A / (1-G_t))^\kappa$, 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 $w_{Ht}/N_t$:

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

(28)

Crucially, as the number of products in the economy increases, the ratio $(V_t^A - V_t^N) / (w_{Ht}/N_t)$ evolves. To see this, we combine (20), (21) and (22) to find the difference in value between an automated and a non-automated firm:

$$V_t^A - V_t^N = \int_t^\infty \exp \left( - \int_t^\tau r u du \right) \left( \pi_t^A - \pi_t^N - \frac{1-\kappa}{\kappa} w_{H\tau} h_{\tau}^A \right) d\tau,$$  

(29)

such that the difference in value between the two types of firms is given by the discounted a result, the free-entry condition would not hold with equality and horizontal innovation would not be sustained. In contrast, the steady-state still exists if the incumbent also automates with positive probability or captures a share of the surplus created by the automation innovation.
difference of the profit flows adjusted for the cost and probability of automation. Further, 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: $w_{Ht}/N_t = \left[ G_t \pi_t^A + (1 - G_t) \pi_t^N \right] / \psi H_t^P$. As a result, the mass of high-skill workers in automation essentially depends on the discounted flow of profits of automated versus non-automated firms divided by the average profits made by firms. Intuitively — from equation (29) — with a positive discount rate, as a first approximation $V_t^A - V_t^N$ will move like $\pi_t^A - \pi_t^N$, so that one gets

$$\frac{V_t^A - V_t^N}{w_{Ht}/N_t} \propto \frac{\pi_t^A - \pi_t^N}{G_t \pi_t^A + (1 - G_t) \pi_t^N} = \frac{1 - (1 + \varphi w_t^{r-1})^{-\mu}}{G_t + (1 - G_t) (1 + \varphi w_t^{r-1})^{-\mu}},$$

(30)

where we used $\pi_t^A - \pi_t^N = \left[ (1 + \varphi w_t^{r-1})^\mu - 1 \right] \pi_t^N$ from (7) and where $\propto$ denotes “approximately proportional”. This highlights low-skill wages (relative to the inverse productivity of machines $\varphi^{-1}$) as the key determinant of automation innovations. Note, that when $w_{Lt} \approx 0$ the incentive for automation innovation is very low, whereas when $w_{Lt} \to \infty$ it approaches $1/G_t > 0$. This price effect bears similarity to Zeira (1998), where the adoption of a labor-saving technology also depends on the price of labor.\footnote{Beyond a focus on different empirical phenomena (an increase in inequality vs. cross-country productivity differences), there are two important differences between our model and Zeira (1998). Zeira (1998) assumes exogenous technological progress (while we model endogenous innovation) and here the innovation cost changes over time while Zeira (1998) has constant adoption cost.}

**Low automation.** When the number of products, $N_t$, is low enough that $w_{Lt}$ is small relative to $\varphi^{-1}$, the difference in profits between automated and non-automated firms is small relative to average profits. Following (28) and (30), the allocation of high-skill labor to automation, $H_t^A$, is low and automation intensity is low. Consequently, as stipulated in Proposition 4.A, 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. If $G_t$ is not initially low, it depreciates following equation (18)). We label this period as Phase 1 (the phases are not qualitatively distinct as automation is low but not exactly 0).

**Rising automation.** As $w_{Lt}$ grows relative to $\varphi^{-1}$, the term $(V_t^A - V_t^N)/(w_{Ht}/N_t)$ increases—more specifically, from (30), it grows approximately proportionally to $(1 + \varphi w_t^{r-1})^\mu - 1$ when $G_t$ is low. This raises the incentive to innovate in automation. Without the externality in the automation technology ($\kappa = 0$), (28) directly implies that $H_t^A$ must rise.
significantly above zero, and with it the Poisson rate of automation, \( \eta \left( \frac{H_t^A}{1 - G_t} \right)^\kappa \) and thereby the share of automated products, \( G_t \). For \( \tilde{\kappa} > 0 \), the depreciation in the share of automated products during Phase 1 gradually makes the automation technology less effective which delays the take-off of automation. It could even potentially prevent the take-off of automation, but our assumption that \( \tilde{\kappa} < 1 - \kappa (1 - \beta) \) rules this out.\(^2\)  

Therefore, during this period—which we refer to as the second phase—the share of automated products \( G_t \) endogenously increases as innovation becomes more directed toward automation. In the context of Figure 2, the relative demand curve pivots counterclockwise and bends upwards while the isocost curve keeps moving outwards. In line with Proposition 1, both the increase in \( G_t \) and \( N_t \) lead to an increase in the skill premium and a decline in the labor share.\(^2\) For some parameters, low-skill wages may even temporarily decline (see numerical examples in Appendix 6.3)—whether they have done so in the data depends on how one accounts for compositional changes in the low-skill population, work benefits and the deflator. Arguably, this time period is the one where our model differs the most from the rest of the literature, in particular 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.\(^2\)

**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 (30) 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 (28)), and so does the Poisson rate of automation, such that \( G_t \) converges to a constant below 1. The economy then converges toward the asymptotic steady-state of Proposition 3 with rising wages and a rising skill premium, in what we label as Phase 3. During this time

\(^2\)Alternatively, if we had assumed that the automation technology obeyed \( \max \{ \eta G_t, \eta \} \left( N_t h_t^A \right)^\kappa \), \( \eta > 0 \), then automation would always take off.

\(^2\)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). As explained below, horizontal innovation declines on average during this time period, which increases \( H_t^P \). In numerical simulations, the net effect is small and changes in \( H_t^P \) do not matter much quantitatively.

\(^2\)At the aggregate level, our model boils down to a nested CES production function (see equation (3)), and Phase 2 corresponds to a period where the share parameter of the composite which features substitutability between machines and low-skill labor, \( G_t \), rises. This change in the share parameter should not be confused with an increase in the elasticity of substitution between machines and low-skill labor (in fact, the Morishima’s elasticity of substitution between these two factors is symmetric and declines in Phase 2 from a value close to \( \epsilon \) to a value close to \( 1 + \beta (\sigma - 1) \)).
period, the share parameters of the nested CES function are constant, but the model continues to differ from a generic capital deepening model since the share parameters are only constant thanks to the interplay between horizontal innovation and automation, and since long-run growth is endogenized and depends on its interaction with automation.

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 our model can reproduce those trends not only qualitatively but also quantitatively.29

3.5 Interactions between automation and horizontal innovation

Before moving to the quantitative exercise, we explore in more details the interactions between automation and horizontal innovation, and show how they can account for some of the 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. Therefore, innovation may be most intensely directed toward automation during the third phase, but the skill premium rising the fastest during the second phase when $G_t$ increases.

29Strictly speaking our model also predicts that automation took off at some point in the past. Of course, automation did not start in the 1970s, which is why we focus on increasing automation instead of its take-off. The goal of our paper is not to deliver an account the long-run economic development since the Industrial Revolution. Yet, our model could be adapted in future research to this context since it delivers long transitional dynamics. In addition, one could easily add a multisector structure to analyze sequential automation in agriculture, manufacturing, etc. In this context, it is worth noting that Katz and Margo (2014) argue that the relative demand for highly skilled workers (in professional, technical and managerial occupations) has increased steadily from perhaps as early as 1820 to the present.
sharply (this is the case both in Section 4 below and in the numerical simulation of Appendix 6.2). 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 \((G_t \text{ and } N_t)\) 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: as automation takes off, horizontal innovation may decline, so that the net effect on growth is ambiguous. Formally, we can show that the rate of horizontal innovation is lower in Phase 3 than in Phase 1 (see proof in Appendix 7.5.5):

**Proposition 5.** 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.

**Effect of innovation parameters.** In Appendix 7.5.6, we establish:

**Proposition 6.** The asymptotic growth rates of GDP \(g^\infty_{GDP}\) and low-skill wages \(g^\infty_w\), increase in the productivity of automation \(\eta\) and horizontal innovation \(\gamma\).

Therefore, in the long-run, a better automation technology (a higher \(\eta\)) actually benefits low-skill wages: the reason is that firms automate faster which encourages horizontal innovation. During the transition, however, a higher \(\eta\) also means that Phase 2 starts sooner, leading to lower low-skill wages at that point and a higher skill premium (see a numerical example in Appendix 7.7.2). These result preview those on the effect of taxes on automation in section 4.4.

---

Proposition 5 establishes that the growth rate of the number of products is higher in a world with no automation at all (as essentially is the case in Phase 1) than in a world with automation, but Proposition 6 shows that conditional on automation happening \((\eta > 0)\), the asymptotic growth rate in the number of products is higher when automation is easier \((\eta\) is higher).
Further results and extensions in the Appendix. i) Our model features elements of self-correction in the presence of exogenous shocks (as Acemoglu and Restrepo, 2017). For instance, with no automation externality ($\bar{\kappa} = 0$), a positive exogenous shock to $G_t$ will be followed by a period where automation is relatively less intense (as the skill premium would decline), so that the asymptotic share of automated products stays the same (see Appendix 7.7.3). ii) Appendix 7.8 studies the social planner’s problem. The optimal allocation is qualitatively similar to the equilibrium we described, which shows that our results are not driven by the market structure we imposed. iii) Appendix 7.9 presents a setting in which automation can only be undertaken before a firm starts production. The transition of the economy is qualitatively identical, which shows that the main results of the paper depend only on the feature that higher $w_{Lt}$ creates incentives to automate and not on the assumption that firms are born non-automated, iv) Appendix 7.10 extends the model to include an endogenous labor supply response.

4 Quantitative Exercise and Policy Experiments

We now conduct a quantitative exercise to compare empirical trends for the United States with the predictions of our model. This allows us to discipline the parameters of our model and subsequently conduct policy experiments.

4.1 Extended model and data

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 path 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:

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

(31)

where $\beta_1 + \beta_2 + \beta_3 = 1$ and $\beta_4 \in [0, 1)$. The central difference between equations (2) and (31) 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 $\Delta$, so that (12) is replaced by:

$$\dot{K}_t = Y_t - C_t - \Delta K_t.$$  \hfill (32)

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 namely $w_{lt}/(w_{ht}^{\beta_2} r_t^{1-\beta_4})$ 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 2A holds with $g_{\infty}^{wh} = g_{\infty}^Y = g_{\infty}^N / ((\beta_2 + \beta_1 \beta_4)(\sigma - 1))$ and $g_{\infty}^{wl} = g_{\infty}^Y (1 + (\sigma - 1) \beta_1 \beta_4) / (1 + \beta_1(\sigma - 1))$. An equivalent to Proposition 3 holds but the system of differential equations includes three control variables and three state variables. Proposition 6 holds as well. The transitional dynamics are similar to that of the baseline model but automation innovation now depends on $w_{lt}/(w_{ht}^{\beta_2} r_t^{1-\beta_4})$. It is low in a first phase, increases in a second phase and stabilizes in the third phase as the economy approaches its asymptotic steady-state. The capital share and the capital output ratio increase in Phase 2 as equipment replaces low-skill labor in production. Details and proofs are provided in Appendix 7.11.

We match our extended model to the data (see Appendix 7.12 for details). Because of data availability and to make our exercise easily comparable to the rest of the literature, 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 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.\footnote{Initial values for $G_t$ and $K_t$ have little impact on the state of the economy several years later.}
**Table 1**: Parameters from quantitative exercise

<table>
<thead>
<tr>
<th>Parameter</th>
<th>σ</th>
<th>ϵ</th>
<th>β₁</th>
<th>γ</th>
<th>˜κ</th>
<th>θ</th>
<th>η</th>
<th>κ</th>
<th>ρ</th>
<th>β₂</th>
<th>Δ</th>
<th>β₄</th>
<th>φ</th>
<th>N₁₉₆₃</th>
<th>G₁₉₆₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>5.96</td>
<td>5.97</td>
<td>0.59</td>
<td>0.45</td>
<td>0.57</td>
<td>0.41</td>
<td>0.65</td>
<td>0.034</td>
<td>0.18</td>
<td>0.011</td>
<td>0.73</td>
<td>1.57</td>
<td>9.9</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

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 1963-2007 (over that time period, the fit is excellent).

### 4.2 Results

The model features a total of 13 parameters with two initial conditions $N_{1923}$ and $G_{1923}$. We will 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 reports 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 5.96, consistent 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 around 6. $\tilde{\kappa}$ is estimated at 0.57 implying a substantial automation externality; a force which causes an accelerated Phase 2. 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 6.4 discusses in details how the parameters are identified.

Figure 3 further shows the predicted path of the matched data series 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 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.
a 5 years moving window, peaks in 1984 at 1.32% and drops to 0.53% in 2004. The total predicted decline in the labor share (8.3 p.p.) is similar to the actual one (6.6 p.p.). 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.

![Graphs showing the predicted and empirical time paths](image)

**Figure 3:** Predicted and empirical time paths

Figure 4 plots 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 3.5 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. In the meantime, the labor share smoothly declines toward its steady-state value of 52.8% and the high-skill labor share increases. Panel C shows that the share of automated products increases sharply through the 1963-2012 time period: In 1963 the value is 2.5% although it is not precisely

---

32Recall 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 recent data show a slow down in software investment (see Beaudry et al., 2016, Eden and Gaggl, 2018).
Figure 4: Transitional Dynamics with calibrated parameters (the growth rates are computed over a 5 year moving average).

identified (see Appendix 6.4) it reaches 20% by 1986 and 64% by 2012 to finally settle at 92% asymptotically. Furthermore, Panel C plots the share of automation innovations among all innovations, \( \eta G_t \left( N_t A_t \right) \kappa_t (1 - G_t) / \left[ \eta G_t \left( N_t h_t A_t \right) \kappa_t (1 - G_t) + G_t g_t N_t \right] \). Even though the share of automated products is very low in the 70s the share of automation innovations is already above 40% in the mid-70s, and it increases steadily till the 2000s. Note, that this ratio peaks in 2005, even though the skill premium is decelerating. As argued in Section 3.5, this occurs in part because the level of automation can be thought as a stock that depreciates with the entry of new products. This constitutes a response to the critique of the literature on skill-biased technical change 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 skill-biased technical change. 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. Finally in the context of Proposition 1, it is worth pointing out that for these parameters, automation is always low-skill labor saving and horizontal innovation low-skill biased.

To assess the predictive power of our model, we reproduce the same exercise but only matching the first 30 years. Appendix 6.4.5 reports the results: the parameters are nearly identical and the calibrated model matches well the rest of the sample period.

Our exercise bears similarities with KORV who also seek to explain the increase in the skill premium using capital-skill complementarity while matching the labor share trend. There are four major differences. First, our exercise is more demanding since instead of directly feeding in the empirical path of equipment, we aim at endogenizing both the
technology path and the equipment stock. 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. This is not an artifact of their specific 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. As a result, as the equipment stock keeps rising (through investment-specific technological change), the income share of equipment declines in the long-run. Fourth, their model does not feature labor-saving innovation as an increase in investment specific technical change increases all wages when capital is perfectly elastic (Appendix 7.13 elaborates on the last two points).

4.3 Data on automation innovations

To match the empirical path on the skill premium and the drop in the labor share, our model requires that the share of automation innovations has increased since the 1970s (see Figure 4.C). We now provide some evidence based on patent data which suggests that this has happened. 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) attempts to classify patents at the US patent office as automation versus non-automation using machine learning techniques (see Appendix 6.5 for details). Figure 5.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 sample), but the increase is of a similar magnitude.

Dechezleprêtre, Hémous, Olsen and Zanella (2019) offer an alternative classification...
system 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 6.5 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 5B. 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.

4.4 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 7.11 for details. In either cases we consider the permanent
unexpected introduction of a 20% tax in the first non-calibrated year, 2013.

Figure 6: 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 [ 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 7.11.5 we show that low-skill wages will increase on impact for any parameter values). 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 $G_t$ (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 $N_t$ (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 $N_t$.

A tax on automation innovation has very different implications: First, high-skill workers move from innovation in automation to production which, on impact, boosts

---

35By comparison in KORV, the effect of such a tax depends on parameters.
36Asymptotically, 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 13 in Appendix 7.11.
output and marginally low-skill wages. As the share of automated products $G_t$ decreases, 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 6 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.

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.\footnote{This feature is also shared by Galor and Weil (2000) and Hansen and Prescott (2002) who endogenize the industrial revolution take-off.}

A lesson from our framework is that if tasks performed by a scarce factor (say labor) can be automated but it is not presently profitable to do so, then, in a growing economy, the return to this factor will eventually increase sufficiently to make it profitable. In other words, there is a long-run tendency for technical progress to displace substitutable labor (a point made by Ray, 2014), but this only occurs if the relevant wages are large relative to the price of machines. This in turn can only happen under three scenarios: either automation itself increases the wages of these workers (the scale effect dominates the substitution effect), or there is another source of technological progress (here, horizontal innovation), or technological progress allows a reduction in the price of machines relative to the consumption good (as in Appendix \ref{app:7.4}). Importantly, when machines are produced with a technology similar to the consumption good, automation can only reduce wages temporarily: a prolonged drop in wages would end the incentives to automate in the first place. Although, we focus on a general equilibrium model with low-skill labor, these insights extend to subsectors of the economy and other scarce factors.

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 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.

References


6 Main Appendix (For Online Publication)

6.1 Formal description of the normalized system of differential equations

We first define an equilibrium as follows:

**Definition 1.** A feasible allocation is defined by time paths of stock of products and share of those that are automated, \([N_t, G_t]_{t=0}^{\infty}\), time paths of use of low-skill labor, high-skill labor, and machines in production \([l_t(i), h_t(i), x_t(i)]_{i \in [0,N_t],t=0}^{\infty}\), a time path of intermediate products \([y_t(i)]_{i \in [0,N_t],t=0}^{\infty}\), time paths of low-skill labor, high-skill labor, and machines in production \([l_t(i), h_t(i), x_t(i)]_{i \in [0,N_t],t=0}^{\infty}\), a time path of intermediate products \([y_t(i)]_{i \in [0,N_t],t=0}^{\infty}\), time paths of high-skill workers engaged in automation \([h_t^A(i)]_{i \in [0,N_t],t=0}^{\infty}\), and in horizontal innovation \([H_t]_{t=0}^{\infty}\), time paths of final good production and consumption levels \([Y_t, C_t]_{t=0}^{\infty}\) such that factor markets clear ((19) holds) and good market clears ((12) holds).

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

Following Proposition 2, asymptotically high-skill wages output and consumption grow proportionately to \(N_t^\psi\) where \(\psi \equiv ((1 - \beta) (\sigma - 1))^{-1}\) when \(g_N^\infty > 0\) and \(G_\infty > 0\) (hence \(\psi\) is the asymptotic elasticity \(Y_t\) with respect to \(N_t\)). Therefore to study the behavior of the system we introduce the normalized variables

\[
\hat{v}_t \equiv w_{H_t} N_t^{-\psi} \quad \text{and} \quad \hat{c}_t \equiv c_t N_t^{-\psi}.
\]

As \(h_t^A\) mechanically tends to 0 as the mass of non-automated firms grows we also introduce

\[
\hat{h}_t^A \equiv N_t h_t^A.
\]

We further define the auxiliary variable

\[
\chi_t \equiv c_t^\theta / \hat{v}_t,
\]

which allows us to simplify the system (\(\chi_t\) is related to the mass of high-skill workers...
in production and therefore, given $\hat{h}_t^A$, to $H^P$ and the growth rate of $N_t$). Since the economy does not feature a non-asymptotic steady-state, we also need to keep track of the level of $N_t$, we do this by introducing

$$n_t \equiv N_t^{-\beta/[(1-\beta)(1+\beta(\sigma-1))]}, \quad (36)$$

which tends toward 0 as $N_t$ tends toward infinity. Finally, we define

$$\omega_t \equiv \left( wLtN_t^{-\psi/(1+\beta(\sigma-1))}\right)^{\beta(1-\sigma)} \quad (37)$$

which asymptotes a finite positive number.

We now derive the system of differential equations satisfied by the normalized variables $(n_t, G_t, h_t, \chi_t)$. (36) immediately gives:

$$\dot{n}_t = -\frac{\beta}{(1-\beta)(1+\beta(\sigma-1))}g_t^N n_t. \quad (38)$$

Rewriting (18) with $\hat{h}_t^A$ gives:

$$\dot{G}_t = \eta G_t^\kappa \left( \hat{h}_t^A \right)^\kappa (1 - G_t) - G_t g_t^N. \quad (39)$$

Defining normalized profits $\hat{\pi}_t^A \equiv N_t^{1-\psi} \hat{\pi}_t^A$ and $\hat{\pi}_t^N \equiv N_t^{1-\psi} \hat{\pi}_t^N$ and the normalized values of firms $\hat{V}_t^A \equiv N_t^{1-\psi} V_t^A$ and $\hat{V}_t^N \equiv N_t^{1-\psi} V_t^N$, then we can rewrite (20) and (21) as

$$\left(r_t - (\psi - 1) g_t^N\right) \hat{V}_t^A = \hat{\pi}_t^A + \hat{V}_t^A, \quad (40)$$

$$\left(r_t - (\psi - 1) g_t^N\right) \hat{V}_t^N = \hat{\pi}_t^N + \eta G_t^\kappa \left( \hat{h}_t^A \right)^\kappa \left( \hat{V}_t^A - \hat{V}_t^N \right) - \hat{\nu}_t \hat{h}_t^A + \hat{V}_t^N. \quad (41)$$

Equation (22) can similarly be rewritten as:

$$\kappa \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa - 1} \left( \hat{V}_t^A - \hat{V}_t^N \right) = \hat{\nu}_t. \quad (42)$$

Equation (23) with equality implies that and $\hat{V}_t^N = \hat{\nu}_t/\gamma$, therefore using (42) into (41), we get:

$$\left(r_t - (\psi - 1) g_t^N\right) \hat{v}_t = \gamma \hat{\pi}_t^N + \gamma \frac{1-\kappa}{\kappa} \hat{\nu}_t \hat{h}_t^A + \hat{V}_t^N. \quad (43)$$

Taking the difference between (40) and (41) and using (42) we obtain:
\[
(r_t - (\psi - 1) g_t^N) \left( \hat{V}_t^A - \hat{V}_t^N \right) = \hat{\pi}_t^A - \hat{\pi}_t^N - \frac{1 - \kappa}{\kappa} \hat{v}_t \hat{h}_t^A + \left( \hat{V}_t^A - \hat{V}_t^N \right).
\]

Using again (42) we get,
\[
(r_t - (\psi - 1) g_t^N) = \kappa \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa-1} \left[ \frac{\hat{\pi}_t^A - \hat{\pi}_t^N}{\hat{v}_t} - \frac{1 - \kappa}{\kappa} \hat{h}_t^A \right] \\
+ \kappa \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa-1} \frac{d}{dt} \left( \frac{\left( \hat{h}_t^A \right)^{1-\kappa}}{\kappa \eta G_t^\kappa} \right) + \hat{v}_t. 
\]

Using (43), we can rewrite this expression as
\[
\gamma \left( \frac{\hat{\pi}_t^N}{\hat{v}_t} + \frac{1 - \kappa}{\kappa} \hat{h}_t^A \right) = \kappa \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa-1} \left[ \frac{\hat{\pi}_t^A - \hat{\pi}_t^N}{\hat{v}_t} - \frac{1 - \kappa}{\kappa} \hat{h}_t^A \right] \\
+ \kappa \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa-1} \frac{d}{dt} \left( \frac{\left( \hat{h}_t^A \right)^{1-\kappa}}{\kappa \eta G_t^\kappa} \right). 
\]

Using (39), this leads to:
\[
\gamma \left( \frac{\hat{\pi}_t^N}{\hat{v}_t} + \frac{1 - \kappa}{\kappa} \hat{h}_t^A \right) = \kappa \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa-1} \left[ \frac{\hat{\pi}_t^A - \hat{\pi}_t^N}{\hat{v}_t} - \frac{1 - \kappa}{\kappa} \hat{h}_t^A \right] \\
+ (1 - \kappa) \frac{\hat{h}_t^A}{\hat{h}_t^A} - \frac{\kappa}{G_t} \left( \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa} \right) \left( 1 - G_t \right) - G_t g_t^N. 
\]

From (36) and (37), we get that \( w_{L,t}^{\beta(1-\sigma)} = \omega_t n_t \), so that, using (7),
\[
\hat{\pi}_t^N = \omega_t n_t \left( \varphi + (\omega_t n_t)^\frac{1}{\sigma} \right)^{-\mu} \hat{\pi}_t^A 
\]
We can then reorder terms in (44) and use (45) to obtain:

\[
\hat{h}_t^A = \frac{\gamma \hat{h}_t^A}{1 - \kappa} \left( \omega t n_t \left( \varphi + (\omega_t n_t)^{\frac{1}{n}} \right)^{-\mu} \frac{\hat{\pi}_t^A}{\hat{v}_t} + \frac{1 - \kappa \hat{h}_t^A}{\kappa} \right) \tag{46}
\]

\[-\frac{\kappa \eta G_t^A}{1 - \kappa} \left( \hat{h}_t^A \right)^{\kappa} \left( 1 - \omega_t n_t \left( \varphi + (\omega_t n_t)^{\frac{1}{n}} \right)^{-\mu} \right) \frac{\hat{\pi}_t^A}{\hat{v}_t} \]

\[+ \eta G_t^A \left( \hat{h}_t^A \right)^{\kappa+1} + \frac{\eta G_t^A}{1 - \kappa} \left( \eta G_t^A - 1 \right) \left( \hat{h}_t^A \right)^{\kappa} \left( 1 - G_t - g_t^N \right) \]

Rewriting (24) using (33), leads to

\[r_t = \rho + \theta \hat{c}_t + \theta \psi g_t^N.\]

Combining this equation with (43) and (45), and using (33), (35) and the definition of \(\hat{\pi}_t^A\) leads to

\[
\dot{\chi}_t = \chi_t \left( \gamma \omega_t n_t \left( \varphi + (\omega_t n_t)^{\frac{1}{n}} \right)^{-\mu} \frac{\hat{\pi}_t^A}{\hat{v}_t} + \frac{1 - \kappa \hat{h}_t^A}{\kappa} - \rho - (\theta \psi - \psi + 1) g_t^N \right). \tag{47}
\]

Together equations (38), (39), (46) and (47) form a system of differential equations which depends on \(\omega_t, \hat{\pi}_t^A / \hat{v}_t\) and \(g_t^N\). To determine \(\hat{\pi}_t^A / \hat{v}_t\), recall that (as proved in the text), profits are given by

\[
\pi(w_L, w_H, \alpha(i)) = \frac{(\sigma - 1)^{\sigma-1}}{\sigma^\sigma} c(w_L, w_H, \alpha(i))^{1-\sigma} Y. \tag{48}
\]

Using (4) and the definition of \(\omega_t\), one gets:

\[
\hat{\pi}_t^A = \frac{(\sigma - 1)^{\sigma-1}}{\sigma^\sigma} \left( \beta^\beta (1 - \beta)^{1-\beta} \right)^{\sigma-1} \left( \varphi + (\omega_t n_t)^{\frac{1}{n}} \right)^{\mu} w_{Ht}^{1-\psi} Y_t. \tag{49}
\]

Rearranging terms in (11) gives

\[
\hat{v}_t = \left( \frac{\sigma - 1}{\sigma} \right)^{\frac{1}{\sigma}} \beta^{\frac{1}{\beta}} (1 - \beta) \left( G \left( \varphi + (\omega_t n_t)^{\frac{1}{n}} \right)^{\mu} + (1 - G) \omega_t n_t \right)^{\psi}. \tag{50}
\]

Using (8), one further gets:

\[Y_t = \sigma \psi \hat{v}_t H_t^P N_t^\psi. \tag{51}\]
Therefore, rewriting (49) with (50) and (51), one gets:

\[
\frac{\hat{\pi}^A_t}{\hat{v}_t} = \frac{\psi \left( \varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu} H_t^P}{G \left( \varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu} + (1 - G) \omega_t n_t},
\]

(52)

which still requires finding \( H_t^P \). Using (4), (5), (6) and aggregating over all automated firms, one gets the following expression for the total demand of machines:

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

Using (50), this expression can be rewritten as:

\[
X_t = \frac{\beta G_t \varphi \left( \varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu-1}}{G \left( \varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu} + (1 - G) \omega_t n_t} Y_t.
\]

(53)

This together with (51) implies that \( \hat{c}_t \) obeys

\[
\hat{c}_t = \left[ 1 - \frac{\sigma - 1}{\sigma} \frac{\beta G_t \varphi \left( \varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu-1}}{G \left( \varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu} + (1 - G) \omega_t n_t} \right]^{\sigma \hat{\psi}_t H_t^P}.
\]

Combining this equation with (35) and (50), leads to

\[
H_t^P = \frac{\left( \frac{\sigma - 1}{\sigma} \right)^{\frac{1}{\sigma} - \frac{\beta}{\sigma}}}{(1 - \beta)^{\frac{\beta}{\sigma} \frac{1}{\sigma} - \frac{\beta}{\sigma} - 1} \chi_t^\frac{1}{\sigma}} \left( \frac{\beta G_t \varphi \left( \varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu-1}}{G \left( \varphi + (\omega_t n_t)^{\frac{1}{\mu}} \right)^{\mu} + (1 - G) \omega_t n_t} \right)^{\psi(\frac{1}{\beta})+1}.
\]

(54)

Using the definition of \( H_t^D \) and (34), one can rewrite (19) for high-skill workers as:

\[
g_t^N = \gamma \left( H - H_t^P - (1 - G_t) \hat{h}_t^A \right).
\]

(55)

Together (52), (54) and (55) determine \( \hat{\pi}_t^A / \hat{v}_t \) and \( g_t^N \) as a function of the original variables \( n_t, G_t, \hat{h}_t^A, \chi_t \) and of \( \omega_t \), which still needs to be determined. To do so, combine
and (11), and use (36) and (37) to obtain an implicit definition of $\omega_t$:

$$
\omega_t = \begin{bmatrix}
\left(\frac{\sigma - 1}{\sigma - \beta}\right)^{1/\gamma} \frac{H^P}{L} \left( G_t \left( \varphi + (\omega t n_t)^{1/\mu} \right)^{1-\mu} + (1 - G_t) \right) \\
\times \left( G_t \left( \varphi + (\omega t n_t)^{1/\mu} \right)^{\mu} + (1 - G_t) \omega t n_t \right)^{\psi - 1}
\end{bmatrix}^{\frac{\beta(1-\sigma)}{1+\beta(\sigma-1)}}. 
$$

(56)

Therefore, the system of differential equations satisfied by $n_t, G_t, \hat{h}^A_t, \chi_t$ is defined by (38), (39), (46) and (47), with $\hat{n}_t^A / \hat{v}_t, H^P_t, g^N_t$ and $\omega_t$ given by (52), (54), (55) and (56). The state variables are $n_t$ and $G_t$ and the control variables $\hat{h}^A_t$ and $\chi_t$. This system admits a steady-state as stipulated below:

**Proposition 7.** Assume that

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

(57)

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

**Proof.** We look for a steady state with positive long-run growth for the system defined by (38), (39), (46) and (47) and we denote such a (potential) steady state $n^*, G^*, \hat{h}^A^*, \chi^*$ (more generally we denote all variables at steady state with an *). Following (38), we immediately get that $n^* = 0$. Using (39), we get that $G^*$ obeys:

$$
G^* = \frac{\eta (G^*)^\kappa (\hat{h}^A^*)^\kappa}{\eta (G^*)^\kappa (\hat{h}^A^*)^\kappa + g^N*}. 
$$

(58)

We focus on a solution with $G^* > 0$ (when $\hat{\kappa} > 0$, $G^* = 0$ is also a solution), this equation implies that with $(g^N)^* > 0$, $G^* < 1$. Then, recalling that $\mu \in (0, 1)$, (56), implies that:

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

Using (52), (47) implies that in steady state,

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

(59)

Note that if $g^N_t = 0$, the economy does not obey this system of equations but that it is also impossible to achieve positive long-run growth, as production is bounded by the production of an economy which has $G_t = 1$. 

45
which uniquely defines \( \hat{h}^{A*} \) as a linear and increasing function of \( g^{N*} \) (recall that \( \theta \geq 1 \)). Note that if \( g^{N*} > 0 \), then \( \hat{h}^{A*} > 0 \).

Then, for \( G^* > 0 \), (58) combined with (59), defines \( G^* \) uniquely as an increasing function of \( g^{N*} \). (55) also uniquely defines \( H^{P*} \) as a function of \( g^{N*} \):

\[
H^{P*} = H - \frac{g^{N*}}{\gamma} - (1 - G^*) \hat{h}^{A*}. \tag{60}
\]

(52) and (58) allows to rewrite (46) in steady state as:

\[
\frac{\eta \kappa (G^*)^{k-1}(\hat{h}^{A*})^k}{1 - \kappa} \psi H^{P*} = \frac{\gamma}{\kappa} (\hat{h}^{A*})^2 + \eta G^t (\hat{h}^{A*})^{k+1}. \tag{61}
\]

Since \( G^*, \hat{h}^{A*} \) and \( H^{P*} \) are functions of \( g^{N*} \), one can rewrite (61) as an equation determining \( g^{N*} \). A steady state with positive growth-rate is a solution to

\[
f\left( g^{N*} \right) \equiv 1 - \frac{\kappa \gamma G^* \hat{h}^{A*}}{\psi H^{P*}} \left( \frac{1}{\kappa} \left( \frac{1}{\hat{h}^{A*}} \right)^{1-\kappa} + \frac{1}{\gamma} \right) = 1, \tag{62}
\]

with \( g^{N*} > 0 \). Indeed, (54) simply determines \( \chi^* \) as:

\[
\chi^* = \left( \frac{\sigma}{\sigma - 1} \right)^{\frac{1}{\theta}} \left( \frac{1 - \beta g^{N*}}{\sigma} \right)^{\frac{1}{\beta}} \left( \frac{1 - \beta}{\beta} \right) \frac{1}{(1 - \beta) \beta (1 - \theta) G^*} = \frac{\theta \sigma}{\beta} \left( H^{P*} \right)^{\frac{\theta}{1 - \theta}} \left( \hat{h}^{A*} \right)^{\frac{1}{1 - \theta}}, \tag{63}
\]

which achieves the characterization of a steady state for the system of differential equations defined by (38), (39), (46) and (47).

In order to establish the sufficiency of equation (25) for positive growth. Note that as \( g^{N*} \to 0 \), then equations (59), (58) and (60) imply that

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

In addition, \( \frac{g^{N*}}{\gamma} + (1 - G^*) \hat{h}^{A*} \) is always greater than \( g^{N*} \gamma \), therefore for a sufficiently large \( g^{N*} \) (smaller than \( \gamma H \)), \( H^{P*} \) is arbitrarily small, while for the same value \( G^* \) and \( \hat{h}^{A*} \) 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 (25).

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 \( G_\infty = G^* \in (0, 1) \), \( g^N = g^{N*} > 0 \) and \( H_\infty^{P} = H^{P*} > 0 \), we can directly use Proposition 2A to obtain the asymptotic growth rates of output and wages (alternatively, we can derive them from the definition of the normalized variables and the fact that those variables are constant in steady-state). \( H^{A*} = (1 - G^*)\hat{h}^{A*} \) is constant in steady-state. Using (50), we have that, as expected, \( \hat{v}_t \) is constant in steady-state, which implies that \( \hat{\pi}^A_t \) is also constant (from (52) and \( n^* = 0 \)). Then using the definition \( \hat{\pi}^A_t \), we obtain that the asymptotic growth rate of profits in automated firms \( \pi^A_t \) is given by (26). We can then use (10) and (15) to obtain the growth rate of profits in non-automated firms (26). Free-entry immediately gives that \( \hat{V}^N_t \) like \( \hat{v}_t \) is constant in steady-state, (42) then implies that \( \hat{V}^A_t \) is also constant in steady-state. We then obtain that the growth rates of the value of automated and non-automated firms obey (27) from the definitions of \( \hat{V}^A_t \) and \( \hat{V}^N_t \). This establishes Proposition 3.

6.2 An illustration of the transitional dynamics

We illustrate our results and further analyze the behavior of our economy through the use of numerical simulations. Unless, otherwise specified, and in line with the results of Sections 2.3 and 3.4, the broad patterns described below do not depend on specific parameter choices and we simply choose “reasonable” parameters (Table 2). Appendix 7.7.4 gives a systematic exploration of the parameter space and section 4 calibrates a richer model to the U.S. data.

Table 2: Baseline Parameter Specification

<table>
<thead>
<tr>
<th>( \sigma )</th>
<th>( \epsilon )</th>
<th>( \beta )</th>
<th>( H )</th>
<th>( L )</th>
<th>( \theta )</th>
<th>( \eta )</th>
<th>( \kappa )</th>
<th>( \hat{\rho} )</th>
<th>( \rho )</th>
<th>( \tilde{\kappa} )</th>
<th>( \gamma )</th>
<th>( N_0 )</th>
<th>( G_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4</td>
<td>2/3</td>
<td>1/3</td>
<td>2/3</td>
<td>2</td>
<td>0.2</td>
<td>0.5</td>
<td>0.25</td>
<td>0.02</td>
<td>0</td>
<td>0.3</td>
<td>1</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**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 1. The initial share of automated products is low, \( G_0 = 0.001 \), but would initially decline.

\[39\text{We employ the so-called “relaxation” algorithm for solving systems of discretized differential equations (Trimborn, Koch and Steger, 2008). See Appendix 7.6 for details.}\]
had we chosen a higher level. We set \( \sigma = 3 \) to capture an initial labor share close to \( 2/3 \). We set \( \tilde{\phi} = 0.25 \) 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 \( (\gamma, \eta, \kappa) \) 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, \( \tilde{\kappa} = 0 \). \( \rho \) and \( \theta \) are chosen such that the interest rate is around 6% initially and asymptotically.

Figure 7 plots the evolution of the economy. Appendix 7.7.1 presents additional graphs on the evolution of wealth and a growth decomposition exercise. 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 \( (V_i^A - V_i^N)/(w_H/N_t) \)—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 \( G_r \) 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 5, 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 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
Figure 7: 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, $(V^A_t - V^N_t) / (w_H/N_t)$, and the skill premium, Panel C the total spending on horizontal innovation and automation as well as the share of automated products (G), and Panel D the wage share of GDP for total wages and low-skill wages.

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 6.3 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 7 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 (13)). 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.
6.3 Negative growth for low-skill wages

This section presents two examples with negative growth for low-skill wages. Appendix 7.7 presents additional numerical results on the transitional dynamics.

![Graphs showing transitional dynamics](image)

**Figure 8**: Transitional Dynamics with temporary decline in low-skill wages with an automation externality. Note: same as for Figure 7 but with an automation externality of $\tilde{\kappa} = 0.49$.

We ensure temporary negative growth in low-skill wages in Figure 8 by setting $\tilde{\kappa} = 0.49$, thereby 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 $(V_t^A - V_t^N) / (w_{Ht}/N_t)$ has already significantly risen (Panel B). Yet, Phase 2 is more intense once it gets started, partly because of the sharp increase in the productivity of the automation technology (following the increase in $G_t$) and partly because low-skill wages are higher. Intense automation puts downward pressure on low-skill wages. At the same time, horizontal innovation drops considerably, both because new firms are less competitive than their automated counterparts, and because the high demand for high-skill workers in automation innovations increases the cost of inventing a new product. This results in a short-lived decline in low-skill wages. Indeed, the decline in $w_{Lt}$ (and increase in high-skill wage $w_{Ht}$) lowers the incentive to automate (Panel B), which in return reduces automation.
Low-skill wages can also drop for $\tilde{\kappa} = 0$ as shown in Figure 9 where low-skill wages slightly decline for a short time period. The associated parameters are given in Table 3. The crucial parameter change is an increase in $\kappa$, 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 3: Baseline Parameter Specification

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>$\epsilon$</th>
<th>$\beta$</th>
<th>$H$</th>
<th>$L$</th>
<th>$\theta$</th>
<th>$\eta$</th>
<th>$\kappa$</th>
<th>$\varphi$</th>
<th>$\rho$</th>
<th>$\tilde{\kappa}$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>73</td>
<td>0.72</td>
<td>0.35</td>
<td>0.65</td>
<td>2</td>
<td>0.2</td>
<td>0.97</td>
<td>0.25</td>
<td>0.022</td>
<td>0</td>
<td>0.28</td>
</tr>
</tbody>
</table>

6.4 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.

6.4.1 Back-of-the-envelope calibration

We first study how the production parameters $\sigma$, $\beta_1$, $\beta_2$, $\beta_4$ and $\Delta$ 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 (168), we get that the skill premium must obey:

\[
\frac{w_{H1963}}{w_{L1963}} \approx \frac{\beta_2}{\beta_1} \frac{L_{1963}}{H_{1963} - \frac{1}{\gamma} g_{N1963}}.
\]

Further, using that most high-skill workers work in production, such that \( \frac{1}{\gamma} g_{N1963} \) is small relative to \( H \), we obtain

\[
\frac{\beta_2}{\beta_1} \approx \frac{w_{H1963} H_{1963}^\gamma}{w_{L1963} L_{1963}^\gamma}, \tag{64}
\]

so that the ratio \( \beta_2/\beta_1 \) 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 \( \beta_2/\beta_1 \). Similarly, using (168), (169), and (175), we get that the labor share in 1963 should obey

\[
ls_{1963} \approx \frac{\beta_2}{\sigma-1} \left( \frac{H}{H_{1963}^\gamma + \beta_1} + \frac{\frac{1}{\gamma} g_{N1963}}{H_{1963}^\gamma + \frac{1}{\gamma} g_{N1963}} \right),
\]

which simplifies into

\[
ls_{1963} \approx \frac{\sigma - 1}{\sigma} (\beta_2 + \beta_1), \tag{65}
\]

if most high-skill workers are in production. Therefore, given \( \sigma \), the initial labor share determines \( \beta_3 \), the ‘external’ capital share. We can then combine (64) and (65) to obtain

\[
\beta_1 \approx \frac{1}{w_{H1963} H_{1963}^\gamma / w_{L1963} L_{1963}^\gamma + 1 - \frac{1}{\sigma}} l_{s1963}, \tag{66}
\]

so that \( \beta_1 \) 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 \( \sigma \) which determines mark-ups.

Combining (178) and (178), 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

\[
g_{2012}^{sp} \approx \frac{\beta_1 (\sigma - 1) (1 - \beta_4)}{1 + \beta_1 (\sigma - 1)} g_{2012}^{GDP} . \tag{67}
\]
Using (168), (169), (175), the labor share now obeys:

\[ l_{s2012} \approx H \left[ \frac{\sigma H}{(\sigma - 1)(\beta_2 + \beta_1 \beta_4)} - \left( \frac{\sigma}{(\sigma - 1)(\beta_2 + \beta_1 \beta_4)} - 1 \right) \left( \frac{1}{H_A^{1963} + 1} \right) \right]^{-1}, \]

which under the assumption that most high-skill workers are in production would simplify again into

\[ l_{s2012} \approx \frac{\sigma - 1}{\sigma} (\beta_2 + \beta_1 \beta_4). \]  \hspace{1cm} (68)

Combining (64), (65) and (68) we obtain:

\[ \beta_4 \approx 1 - \left( 1 - \frac{l_{s2012}}{l_{s1963}} \right) \left( \frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}} + 1 \right). \]  \hspace{1cm} (69)

Therefore, in this approximation, \( \beta_4 \) 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 \( l_{s2012} \) by the lowest value over 1963-2012 (which is 57%) and \( l_{s1963} \) 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 \( l_{s2012} \) is higher than the asymptotic value of the labor share).

Using (66), (67) and (69) we obtain:

\[ \sigma \approx \frac{1}{l_{s1963} \left[ \frac{\sigma_{GDP}^{1963}}{\sigma_{GDP}^{2012}} \left( 1 - \frac{l_{s2012}}{l_{s1963}} \right) - \frac{1}{\frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}} + 1} \right]}. \]

that is given the initial wage bill ratio and the labor shares in 1963 and 2012, which inform us about \( \beta_1, \beta_2 \) and \( \beta_4, \sigma \) is determined by the ratio between the growth rate of GDP and that of the skill-premium in the third phase. The larger is \( \sigma \), the more automated firms gain over non-automated ones and therefore the more the skill premium rises relative to GDP: hence a lower \( \frac{\sigma_{GDP}^{1963}}{\sigma_{GDP}^{2012}} \) is associated with a larger \( \sigma \). When using the last 10 years to determine \( \frac{\sigma_{GDP}^{1963}}{\sigma_{GDP}^{2012}} \), we find that \( \sigma \approx 6.77 \), while our estimation procedure leads to \( \sigma = 5.96 \).

Given \( \sigma \), one can then find \( \beta_1 \) using (66), 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 \( \sigma \) provided that \( \sigma \) is large enough). Using (64), we then obtain \( \beta_2 \approx 0.28 \) which is higher than the
estimated value of 0.18, in line with the fact that [64] gives an overestimate of $\beta_2/\beta_1$.

To get a proxy for $\Delta$, we look at the steady-state value for the equipment to GDP ratio. Using (194), (195) and the definition of GDP, we obtain that

$$\frac{\hat{K}^*}{GDP} = \frac{1}{\tilde{r}^*} \frac{\beta_3 + \beta_1 (1 - \beta_4)}{\sigma - 1 + (\beta_2 + \beta_1 \beta_4) \left( \frac{H}{H^{ss}} - 1 \right)}.$$ 

Denote by $K_{eq}$ the stock of capital used as equipment, we get

$$\hat{K}_{eq}^* = \frac{\beta_1 (1 - \beta_4)}{\beta_3 + \beta_1 (1 - \beta_4)} \hat{K}^*,$$

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 (193), we obtain

$$\frac{\hat{K}_{eq}^*}{GDP} = \frac{1}{\rho + \Delta + \theta g_{GDP}^*} \frac{\beta_1 (1 - \beta_4)}{\sigma - 1 + (\beta_2 + \beta_1 \beta_4) \left( \frac{H}{H^{ss}} - 1 \right)}.$$ 

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

$$\rho + \Delta + \theta g_{2000} = \left( \frac{\sigma - 1}{\sigma} \right) \frac{\beta_1 (1 - \beta_4)}{K_{eq,2000}/GDP_{2000}} \approx 0.051$$

using the values computed above. It is therefore not surprising that we find a low $\Delta$ in the estimation. This is due in particular to the high-level of $K_{eq,2000}/GDP_{2000} = 1.5$ (with the actual estimated values for $\sigma$, $\beta_1$ and $\beta_4$ we would still find that $\rho + \Delta + \theta g_{2000}^{GDP} \approx 0.088$).

### 6.4.2 Parameters’ effect on empirical moments

In this section, we illustrate 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 4 reports the elasticities (note that $\beta_3$ is completely determined by $\beta_1$ and $\beta_2$).

The initial skill premium is most strongly affected by the production function pa-
rameters $\beta_1, \beta_2$ and $\beta_4$: A higher share of high-skill workers in production, $\beta_2$, directly increases the skill-premium. A higher value of $\beta_4$ makes automation more expensive, which increases the demand for low-skill workers and reduces the skill premium. A higher $\beta_1$ implies a lower $\beta_3$ 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. $\beta_2$ has the opposite effect on the skill premium in 2012. A higher $\beta_2$ reduces the multiplier of $N_t$ on output, $Y_t$ which reduces the growth rate of the economy. The automation technology parameters, $\kappa, \tilde{\kappa}, \eta$ also have a large effect on the skill premium in 2012.

The initial labor share depends on $\beta_1, \beta_2$ and $\beta_4$, 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 $\sigma$ since we keep the stock of products in 1963 constant. Both $\beta_1$ and $\beta_2$ 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 $\gamma, \eta$ lead to higher growth and therefore higher $GDP/labor$ in 2012, though naturally not in 1963.

Capital equipment / GDP in 1963 depends positively on $\beta_1$ and negatively on $\beta_4$ because the initial capital stock is fixed. For 2012, a higher $\beta_4$ increases the cost of automation and thereby reduces $K_{eq}/GDP$. Horizontal innovation productivity, $\gamma$, 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, $\tilde{\phi}$, shifts capital into equipment and consequently raises $K_{eq}/GDP$.

6.4.3 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 $\theta_i$, consider

$$V(\theta_i, \bar{\theta}_{-i}(\theta_i)),$$
Increases in $\rho, \theta, \eta, \gamma$ all govern the growth rate of the economy and are weakly identified individually. Proposition 2 makes clear, does not govern the asymptotic growth of income inequality.

Parameters Skill premium Labor share GDP/labor $K_{eq}/GDP$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$</td>
<td>-0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>-0.5</td>
<td>-1.9</td>
<td>2.1</td>
<td>1.1</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.2</td>
<td>-0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.3</td>
<td>-0.1</td>
<td>0.7</td>
<td>0.6</td>
<td>-1.0</td>
<td>-2.6</td>
<td>7.7</td>
<td>1.8</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.1</td>
<td>-0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.7</td>
<td>-0.2</td>
<td>-0.5</td>
</tr>
<tr>
<td>$\bar{\kappa}$</td>
<td>-0.2</td>
<td>-0.2</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
<td>-0.4</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.0</td>
<td>-0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.3</td>
<td>0.1</td>
<td>-0.2</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.1</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>-0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>-0.1</td>
<td>-0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.3</td>
<td>0.3</td>
<td>-0.2</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.1</td>
<td>-0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.6</td>
<td>0.2</td>
<td>-0.3</td>
</tr>
<tr>
<td>$\bar{\varphi}$</td>
<td>0.1</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.4</td>
<td>-0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>-0.6</td>
<td>-1.5</td>
<td>1.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>$\Delta$</td>
<td>0.0</td>
<td>-0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.1</td>
<td>0.0</td>
<td>-0.2</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.7</td>
<td>-1.9</td>
<td>0.1</td>
<td>0.6</td>
<td>-0.1</td>
<td>-1.3</td>
<td>-5.0</td>
<td>-3.7</td>
</tr>
<tr>
<td>$N_{1963}$</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.3</td>
<td>-0.2</td>
<td>0.0</td>
</tr>
<tr>
<td>$G_{1963}$</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-0.1</td>
<td>0.6</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 4: The effect of parameters on the four empirical paths (numbers refer to elasticities of empirical value wrt. parameter)

where $\bar{\theta}_i(\theta_i)$ are the parameters that minimize $V$ for any given $\theta_i$ and $\bar{\theta}_i = \arg\min_{\theta_i} V(\theta_i, \bar{\theta}_i(\theta_i))$ is the minimizing value of $\theta_i$. Consequently, a Taylor expansion around $\bar{\theta}_i$ yields:

$$\frac{V(\theta_i, \bar{\theta}_i(\theta_i)) - V(\bar{\theta}_i, \bar{\theta}_i(\theta_i))}{V(\theta_i, \bar{\theta}_i(\theta_i))} \approx \left(\frac{\theta_i - \bar{\theta}_i}{\bar{\theta}_i}\right)^2 \approx \frac{1}{2V(\bar{\theta}_i, \bar{\theta}_i(\theta_i))} \frac{d^2V(\bar{\theta}_i, \bar{\theta}_i(\theta_i))}{d\theta_i^2} \bar{\theta}_i^2.$$

We compute the expression on the left. The results are in Table 5 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 $\epsilon$, 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 $\varphi$ can to a certain extent be accommodated by changes to $N_{1963}$ and consequently neither is very well-identified. The depreciation of capital $\Delta$ is also not well identified because it mostly affects the growth rate of the capital stock which also depends on $\rho$ and $\theta$ (equation [70]). Given that this parameter is the one estimated outside a common range this is a reassuring finding.
Table 5: The “curvature” of deviating from the optimal parameter

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\sigma$</th>
<th>$\epsilon$</th>
<th>$\beta_1$</th>
<th>$\gamma$</th>
<th>$\tilde{\kappa}$</th>
<th>$\theta$</th>
<th>$\eta$</th>
<th>$\kappa$</th>
<th>$\rho$</th>
<th>$\tilde{\varphi}$</th>
<th>$\beta_2$</th>
<th>$\Delta$</th>
<th>$\beta_4$</th>
<th>$N_{1963}$</th>
<th>$G_{1963}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curvature</td>
<td>82.0</td>
<td>2.7</td>
<td>219.3</td>
<td>3.7</td>
<td>1.7</td>
<td>4.2</td>
<td>3.0</td>
<td>47.1</td>
<td>4.8</td>
<td>4.0</td>
<td>12.8</td>
<td>2.4</td>
<td>2142.1</td>
<td>2.3</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Figure 10: Predicted and empirical time paths for $\tilde{\kappa} = 0$

6.4.4 The role of the automation externality

To analyze more specifically the role played by the automation externality, we recalibrate our model but without the automation externality (i.e., we impose that $\tilde{\kappa} = 0$). Figure 10 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 require an “accelerated Phase 2” 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 5).

6.4.5 Out-of-sample prediction

Finally, we reproduce our calibration exercise but only trying to match the first 30 years of data. Figure 11 reports the results and Table 6 gives the new parameters. The model
behaves very well out of sample. The predicted path and parameters are close to those of the baseline regression (compare with Table 1). The calibrated model over the first 30 years slightly underestimates the pace of the rise of the skill-premium. The biggest difference is that the elasticity of substitution between low-skill workers and machines is estimated considerably lower at 4.3 instead of 6. This creates less of an incentive for automation slows the growth in the skill-premium and lowers the growth rate of equipment (Panel D).

### 6.5 Patent data on automation innovations

The data for Figure 5.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

| Parameter | Value | \( \sigma \) | 4.34 | \( \epsilon \) | 0.60 | \( \beta_1 \) | 0.43 | \( \gamma \) | 0.55 | \( \hat{\kappa} \) | 1 | \( \theta \) | 0.40 | \( \eta \) | 0.66 | \( \kappa \) | 0.034 | \( \rho \) | 0.16 | \( \Delta \) | 0.011 | \( \beta_4 \) | 0.76 | \( \tilde{\varphi} \) | 1.56 | \( N_{1963} \) | 14.8 | \( G_{1963} \) | 0.03 |
|-----------|-------|-------------|-----|-------------|----|-------------|----|-------------|----|-------------|----|-------------|----|-------------|----|-------------|----|-------------|----|-------------|----|-------------|----|-------------|----|-------------|----|

**Table 6:** Parameters (only matching the first 30 years)
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 5.B, we use granted patents at the USPTO.

7 Secondary Appendix (For Online Publication)

7.1 Relationship between wages and \( N \) and \( G \)

7.1.1 Imperfect substitute case: \( \epsilon < \infty \).

We first focus on the imperfect substitute case. Rewrite (10) as

\[
\frac{w_H}{w_L} = \frac{1 - \beta}{\beta} \frac{L}{H^P} \frac{G (1 + \varphi w_L^{-1})^{\mu} + 1 - G}{G (1 + \varphi w_L^{-1})^{\mu-1} + 1 - G}.
\]

(71)

Since \( 0 < \mu < 1 \), (71) establishes \( w_H \) as a function of \( G, H^P \) and \( w_L \) (but not \( N \)) 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 \( G > 0 \). (11) similarly establishes \( w_H \) as a function of \( N, G \) and \( w_L \) (but not \( H^P \)), \( w_H \) is decreasing in \( w_L \) and increasing in \( N \) and \( G \). It is then immediate that \( w_H, w_L \) are jointly uniquely determined by (71) and (11) for given \( N, G \) and \( H^P \), both increase in \( N \), and \( w_H \) increases in \( G \) (in addition, (11) traces an iso-cost curve in the input prices...
plan which is convex).

In addition, (71) 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 (and following (13) the labor share decreases in \( N \)). Assume that \( w_L \) decreases in \( G \), then since \( w_H \) increases in \( G \), we immediately get that \( w_H/w_L \) increases in \( G \). Assumnes now that \( w_L \) increases in \( G \), then the right-hand side of (71) increases with \( G \) both directly and because \( w_L \) increases, this ensures that \( w_H/w_L \) increases in \( G \) (and following (13) the labor share decreases in \( G \)). Therefore both an increase in \( N \) and an increase in \( G \) are skill-biased.

**Comparative statics of \( w_L \) with respect to \( G \).** We now analyze how \( w_L \) changes with \( G \) (for given \( N \) and \( H^P \)). To do so, we combine both equations to get:

\[
\hat{w}_L = \left( \frac{\sigma-1}{\sigma} \beta \left( \frac{H^P}{L} \right)^{(1-\beta)} N^{\frac{1}{1-\beta}} \left( G \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{1-\mu} + (1-G) \right)^{1-\beta} \times \left( G \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu} + (1-G) \right)^{-\frac{1}{1-\beta}} \right) \frac{\hat{G}}{\text{Den}},
\]

(72)

Log differentiating with respect to \( G \) one obtains:

\[
\frac{\hat{w}_L}{\hat{G}} = \left[ \frac{1}{\sigma-1} G \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu} + (1-G) \right] - (1-\beta) \left[ \frac{1 - \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu-1}}{G \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu-1} + (1-G)} + \frac{\left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu-1} - 1}{G \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu} + (1-G)} \right]
\]

(73)

where

\[
\text{Den} \equiv 1 - \frac{\beta \varphi w_{L}^{\epsilon-1} \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu}}{\left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu} + (1-G)} + \frac{\varphi w_{L}^{\epsilon-1} (\epsilon-1) (1-\beta) \left( \mu G \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu} + (1-G) \right)^{1-\mu} + (1-\mu) G \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu-1}}{G \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu} + (1-G)}.
\]

\( \text{Den} > 0 \) as \( \epsilon > 1 \), \( \mu \in (0,1) \) and \( \frac{\beta \varphi w_{L}^{\epsilon-1} \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu}}{1 + \varphi w_{L}^{\epsilon-1} \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu} + (1-G)} < 1 \). In (73) the scale effect term is positive as \( \left( 1 + \varphi w_{L}^{\epsilon-1} \right)^{\mu-1} - 1 > 0 \). This term comes from the differentiation of (11) 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 (10) 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 on, the other hand \( \frac{1}{\sigma - 1} > 1 - \beta \), then the scale effect is dominated by the substitution effect provided that \( \frac{1 - (1 + \varphi w_L^{-1})^{\mu - 1}}{G(1 + \varphi w_L^{-1})^{\mu - 1} + 1 - G} \) is large enough. From (72) we get:

\[
\begin{align*}
w_L &= \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_P}{L} \right)^{(1 - \beta)} N^{\frac{1}{\sigma - 1}} \left( 1 - G \frac{1 - (1 + \varphi w_L^{-1})^{\mu - 1}}{G(1 + \varphi w_L^{-1})^{\mu - 1} + 1 - G} \right)^{1 - \beta} (G ((1 + \varphi w_L^{-1})^{\mu} - 1) + 1)^{\frac{1}{\sigma - 1}} \\
&> \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_P}{L} \right)^{(1 - \beta)} N^{\frac{1}{\sigma - 1}} \left( (1 + \varphi w_L^{-1})^{-1} \right)^{1 - \beta},
\end{align*}
\]

where the last line uses that \( G \in [0,1] \). We then obtain that

\[
w_L \left( 1 + \varphi w_L^{-1} \right)^{1 - \beta} > \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_P}{L} \right)^{(1 - \beta)} N^{\frac{1}{\sigma - 1}},
\]

which ensures that \( \lim_{N \to \infty} w_L = \infty \) uniformly with respect to \( G \) (i.e. for any \( \bar{w}_L > 0 \), there exist \( N \) such that for any \( N > \bar{N} \) and any \( G, w > \bar{w}_L \)). Since

\[
\lim_{w_L \to \infty, G \to 1} \frac{1 - (1 + \varphi w_L^{-1})^{\mu - 1}}{G(1 + \varphi w_L^{-1})^{\mu - 1} + 1 - G} = \infty,
\]

we get that

\[
\lim_{N \to \infty, G \to 1} \frac{1 - (1 + \varphi w_L^{-1})^{\mu - 1}}{G(1 + \varphi w_L^{-1})^{\mu - 1} + (1 - G)} = \infty.
\]

Therefore for \( N \) and \( G \) large enough the substitution effect dominates. This achieves the proof of Proposition 1.

**Increase in the number of non-automated products.** To study the effect of an increase in the number of non-automated products only, we log differentiate (72) with...
respect to both $N$ and $G$ and obtain:

$$\hat{\omega}_L = \left[ \left( \frac{1}{\sigma - 1} - (1 - \beta)\left(1 + \omega_L^{-1}\right)^{\mu - 1} \right) \frac{G(1 + \omega_L^{-1})^{\mu} + (1 - G)}{G(1 + \omega_L^{-1})^{\mu - 1} + (1 - G)} \right] \hat{G} + \frac{1}{\sigma - 1} \hat{\tilde{N}} \right] \frac{1}{\text{Den}}$$

In that case $NG$ is a constant, so that $\hat{G} = -\hat{N}$, therefore denoting $\hat{\omega}_L^{NT}$ ($NT$ for “new tasks”), the change in $w_L$, we get:

$$\hat{\omega}_L^{NT} = \left[ \left( \frac{1 - \beta}{G(1 + \omega_L^{-1})^{\mu} + (1 - G)} \right) + \frac{(1 - \beta)G\left(1 - (1 + \omega_L^{-1})^{\mu - 1}\right)}{G(1 + \omega_L^{-1})^{\mu - 1} + (1 - G)} \right] \hat{N} \frac{1}{\text{Den}}.$$  (74)

Hence low-skill wages always increase with the arrival of non-automated products. Log-differentiating [10], one gets:

$$\hat{\omega}_H - \hat{\omega}_L = \left( \frac{\mu G (1 + \varphi w_L^{-1})^\mu}{G (1 + \varphi w_L^{-1})^\mu + 1 - G} + \frac{(1 - \mu) G (1 + \varphi w_L^{-1})^{\mu - 1}}{G (1 + \varphi w_L^{-1})^{\mu - 1} + 1 - G} \right) \frac{\epsilon - 1}{1 + \varphi w_L^{-1}} \hat{\omega}_L$$

$$+ \left( \frac{G (1 + \varphi w_L^{-1})^\mu - 1}{G (1 + \varphi w_L^{-1})^\mu + 1 - G} + \frac{G (1 - (1 + \varphi w_L^{-1})^{\mu - 1})}{G (1 + \varphi w_L^{-1})^{\mu - 1} + 1 - G} \right) \hat{G}.$$  

Using (74) and that $\hat{G} = -\hat{\tilde{N}}$, then we get that following a change in the mass of non-automated products (keeping the mass of automated products constant):

$$\hat{\omega}_H^{NT} = \left[ \left( \frac{(1 - \mu) G (1 + \varphi w_L^{-1})^{\mu - 1}}{G (1 + \varphi w_L^{-1})^{\mu - 1} + 1 - G} \right) \frac{\epsilon - 1}{1 + \varphi w_L^{-1}} + 1 \right]$$

$$\times \left[ \frac{G (1 + \varphi w_L^{-1})^\mu - 1}{G (1 + \varphi w_L^{-1})^\mu + 1 - G} + \frac{G (1 - (1 + \varphi w_L^{-1})^{\mu - 1})}{G (1 + \varphi w_L^{-1})^{\mu - 1} + 1 - G} \right] \hat{\tilde{N}} \frac{1}{\text{Den}}$$

$$- \left( \frac{G (1 + \varphi w_L^{-1})^\mu - 1}{G (1 + \varphi w_L^{-1})^\mu + 1 - G} + \frac{G (1 - (1 + \varphi w_L^{-1})^{\mu - 1})}{G (1 + \varphi w_L^{-1})^{\mu - 1} + 1 - G} \right) \hat{N}. $$

We then obtain:
Therefore an increase in the mass of non-automated products leads to higher high-skill wages. Finally we obtain

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

$$
= \frac{(\sigma-1)^{-1}}{G(1+\varphi w_L^{-1})^\mu+1-G} \left( 1 + \frac{G(1-\mu)(1+\varphi w_L^{-1})^\mu-1}{G(1+\varphi w_L^{-1})^\mu+1-G} (\epsilon - 1) \varphi w_L^{-1} \right) \frac{\hat{N}}{\text{Den}}
$$

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

### 7.1.2 Perfect substitute case: $\epsilon = \infty$

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

$$
w_L = \frac{\sigma - 1}{\sigma} \beta \left( \frac{H^P}{L} \right)^{1-\beta} N_{\sigma-1} \frac{1}{\sigma-1}, \quad (75)
$$

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

Case ii) $w_L = \tilde{\varphi}^{-1}$: automated firms use machines but also possibly workers, in which case high-skill wages can be obtained from (11) which is now written as:

$$
\frac{\sigma}{\sigma - 1} \beta \left( \frac{1+\varphi w_L^{-1}}{\sigma-1} \right)^{1-\beta} \frac{\tilde{\varphi}^{-1} w_H^{1-\beta}}{\tilde{\varphi}^{-1}} = 1.
$$
Case iii) \( w_L > \tilde{\varphi}^{-1} \) and all automated firms use machines only, in that case, we get that (72) is replaced by

\[
\begin{align*}
\frac{w_L}{\sigma - 1} \beta \left( \frac{H \phi}{L} \right)^{(1-\beta)} N^{\frac{1}{\beta-1}} (1 - G)^{1-\beta} \left( G \left( \tilde{\varphi} w_L \right)^{\beta(\sigma-1)} + 1 - G \right)^{\frac{1}{\beta-1} - (1-\beta)},
\end{align*}
\]

and the skill premium obeys:

\[
\begin{align*}
\frac{w_H}{w_L} = \frac{1 - \beta}{\beta} \frac{L}{H P} \frac{G (w_L \tilde{\varphi})^{\beta(\sigma-1)} + 1 - G}{1 - G}.
\end{align*}
\]

One can rewrite (76) as

\[
\begin{align*}
\frac{w_H}{w_L} = \frac{1 - \beta}{\beta} \frac{L}{H P} \frac{G (w_L \tilde{\varphi})^{\beta(\sigma-1)} + 1 - G}{1 - G}.
\end{align*}
\]

The left-hand side increases in \( w_L \) and the right-hand side decreases in \( w_L \), hence this expression defines \( w_L \) uniquely, in addition, the solution is greater than \( \tilde{\varphi}^{-1} \) if and only if

\[
\begin{align*}
N^{\frac{1}{\beta-1}} (1 - G)^{1-\beta} > \frac{\sigma}{(\sigma - 1) \beta \tilde{\varphi}} \left( \frac{L}{H P} \right)^{(1-\beta)}.
\end{align*}
\]

Hence \( w_L \) and \( w_H \) are defined uniquely as functions of \( N, G \) and \( H P \). If \( N^{\frac{1}{\beta-1}} < \frac{\sigma}{(\sigma - 1) \beta \tilde{\varphi}} \left( \frac{L}{H P} \right)^{(1-\beta)} \), we are in case i), if \( N^{\frac{1}{\beta-1}} (1 - G)^{1-\beta} \leq \frac{\sigma}{(\sigma - 1) \beta \tilde{\varphi}} \left( \frac{L}{H P} \right)^{(1-\beta)} \leq N^{\frac{1}{\beta-1}} \) then we are in case ii) and if \( N^{\frac{1}{\beta-1}} (1 - G)^{1-\beta} > \frac{\sigma}{(\sigma - 1) \beta \tilde{\varphi}} \left( \frac{L}{H P} \right)^{(1-\beta)} \), 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, (76) 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 (76), one gets:

\[
\begin{align*}
\hat{w}_L = \left[ \left( \frac{1}{\sigma - 1} - (1 - \beta) \right) \left( (\tilde{\varphi} w_L)^{\beta(\sigma-1)} - 1 \right) G \right] \frac{1 - \beta}{\beta} \frac{G}{1 - G} \left[ \tilde{G} \right].
\end{align*}
\]
where
\[
\text{Den} \equiv 1 - \beta \frac{G (\hat{\varphi}_wL)^{\beta(\sigma-1)}}{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} + 1 - G} + \frac{(1 - \beta) \beta (\sigma - 1) G (\hat{\varphi}_wL)^{\beta(\sigma-1)}}{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} + 1 - G}.
\]  
(79)

We have
\[
\left( \frac{1}{\sigma - 1} - (1 - \beta) \right) \frac{((\hat{\varphi}_wL)^{\beta(\sigma-1)} - 1)}{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} + 1 - G} - (1 - \beta) \frac{G}{1 - G} < \frac{1}{\sigma - 1} - \frac{1 - \beta}{1 - G},
\]
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 (76) with respect to both \( N \) and \( G \), one gets:
\[
\hat{w}_L = \left[ \left( \frac{1}{\sigma - 1} - (1 - \beta) \right) \frac{((\hat{\varphi}_wL)^{\beta(\sigma-1)} - 1)}{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} + 1 - G} - (1 - \beta) \frac{G}{1 - G} \right] \hat{G} + \frac{1}{\sigma - 1} \hat{N} \left\{ \frac{1}{\text{Den}} \right. 
\]

For an increase in the mass of non-automated products, \( \hat{G} = -\hat{N} \), so that the change in \( w_L \) in that case is given by:
\[
\hat{w}_L^{NT} = \left(1 - \beta\right) \left( \frac{((\hat{\varphi}_wL)^{\beta(\sigma-1)} - 1)}{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} + 1 - G} + \frac{G}{1 - G} \right) + \frac{1}{\sigma - 1} \left(1 - \frac{((\hat{\varphi}_wL)^{\beta(\sigma-1)} - 1)}{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} + 1 - G} \right) \hat{N} \left\{ \frac{1}{\text{Den}} \right. 
\]
Therefore \( w_L \) increases.

Log-differentiating (77), we get:
\[
\hat{w}_H = \frac{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} - 1}{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} + 1 - G} \hat{G} + \frac{G}{1 - G} \hat{G} + \left(1 + \frac{\beta (\sigma - 1) G (\hat{\varphi}_wL)^{\beta(\sigma-1)}}{G (\hat{\varphi}_wL)^{\beta(\sigma-1)} + 1 - G} \right) \hat{w}_L.
\]
Therefore, for an increase in the mass of non-automated products, one gets:

\[
\hat{w}_{NT}^H = \frac{1}{\sigma - 1} \left( 1 - \left( \bar{\varphi} w_L^{\beta(\sigma - 1)} - 1 \right) \frac{G}{G (\bar{\varphi} w_L^{\beta(\sigma - 1)} + 1 - G) - 1} \right) \frac{\hat{N}}{\text{Den}},
\]

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

\[
\hat{w}_{NT}^H - \hat{w}_{NT}^L = - \frac{(1 - \beta) G (\bar{\varphi} w_L^{\beta(\sigma - 1)})}{\left( G (\bar{\varphi} w_L^{\beta(\sigma - 1)} + 1 - G) (1 - G) \right)} \frac{\hat{N}}{\text{Den}}.
\]

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)).

**Comparison with an increase in machine’s productivity.** We now look at the effect of an increase in machine’s productivity \( \bar{\varphi} \) (which up to some relabeling is equivalent to a decline in the price of machines). We focus on the case \( \epsilon < 1 \), so that an increase in \( \bar{\varphi} \) is equivalent to an increase in \( \varphi \). To look at its effect on low-skill wages, log-differentiate (72)

\[
\hat{w}_L = \frac{G w_L^{\epsilon - 1} (1 + \varphi w_L^{\epsilon - 1})^{\mu - 1}}{\text{Den} (1 + \varphi w_L^{\epsilon - 1})} \left( \frac{(1 - \beta) (\mu - 1)}{G (1 + \varphi w_L^{\epsilon - 1})^{\mu - 1} + 1 - G} + \frac{\mu (\frac{1}{\sigma - 1} - (1 - \beta)) (1 + \varphi w_L^{\epsilon - 1})}{G (1 + \varphi w_L^{\epsilon - 1})^{\mu} + 1 - G} \right) \hat{\varphi},
\]

where \( \text{Den} \) is still given by (79). 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

\[
\frac{\partial w_L}{\partial \varphi} < 0 \iff \frac{G (1 + \varphi w_L^{\epsilon - 1})^{\mu} + 1 - G}{G (1 + \varphi w_L^{\epsilon - 1})^{\mu - 1} + 1 - G \psi - 1} \frac{1}{\mu (1 + \varphi w_L^{\epsilon - 1})} > \frac{(1 - \beta) (\mu - 1)}{1 - \mu}.
\]

Using (73), we get

\[
\frac{\partial w_L}{\partial G} < 0 \iff \frac{G (1 + \varphi w_L^{\epsilon - 1})^{\mu} + 1 - G}{G (1 + \varphi w_L^{\epsilon - 1})^{\mu - 1} + 1 - G \psi - 1} \frac{1}{\mu (1 + \varphi w_L^{\epsilon - 1})} > \frac{(1 + \varphi w_L^{\epsilon - 1})^{\mu - 1}}{1 - (1 + \varphi w_L^{\epsilon - 1})^{\mu - 1}}.
\]
Note that
\[
\frac{(1 + \varphi w_{L}^{\varepsilon-1})^{\mu} - 1}{1 - (1 + \varphi w_{L}^{\varepsilon-1})^{\mu-1}} < \frac{\mu (1 + \varphi w_{L}^{\varepsilon-1})}{1 - \mu} \iff (1 + \varphi w_{L}^{\varepsilon-1})^{\mu} < 1 + \mu \varphi w_{L}^{\varepsilon-1},
\]
which is always true since \( \mu < 1 \). Therefore, \( \frac{\partial w}{\partial \varepsilon} < 0 \) implies that \( \frac{\partial w}{\partial \mu} < 0 \) but the reverse is not true.

7.2 Proofs of the asymptotic results

7.2.1 Proof of Proposition 2

Case with \( G_{\infty} > 0 \) (Parts A and B). To see that \( w_{Lt} \) is bounded from below, assume that \( \liminf w_{Lt} = 0 \). Then using that \( H_{t}^{P} \) and \( G_{t} \) admit positive limits, \( (10) \) implies that \( \liminf w_{Ht} = 0 \). Plugging this further in \( (11) \) gives \( \liminf N_{t} = 0 \), which is impossible since \( g_{N_{t}} \) admits a positive limit. Therefore, \( w_{Lt} \) must be bounded below, so that \( (11) \) gives \( g_{w} H_{\infty} = \psi g_{N_{\infty}} \). Further, using that \( H_{t}^{P} \) admits a limit and \( (8) \) gives the growth rate of \( Y_{t} \). We now derive the asymptotic growth rate of \( w_{Lt} \). To do so we consider in turn the case where \( \varepsilon < \infty \), and the case where \( \varepsilon = \infty \).

Subcase with \( \varepsilon < \infty \). We use equation \( (72) \) which gives \( w_{Lt} \) as a function of \( N_{t}, G_{t} \) and \( H_{t}^{P} \). Note that assuming that \( w_{Lt} \) is bounded above leads to a contradiction, therefore \( \lim w_{Lt} = \infty \).

Assume first that \( G_{\infty} < 1 \), then, since \( \lim w_{Lt} = \infty \), \( (72) \) implies
\[
w_{Lt} \sim \left( \frac{\sigma - 1}{\sigma} \beta \right)^{\frac{1}{\beta+1}} (1 - G_{\infty}) \frac{H_{t}^{P}}{L} (\varphi^{\mu})^{\psi-1} \right)^{\frac{1}{1+\beta(\sigma-1)}} N_{t}^{\frac{\psi}{1+\beta(\sigma-1)}},
\]
where for \( x_{t} \) and \( y_{t} \) (possibly with no limits), \( x_{t} \sim y_{t} \) signifies \( x_{t}/y_{t} \to 1 \). This delivers Part A).

Consider now the case where \( G_{\infty} = 1 \). Note that \( (72) \) gives:
\[
w_{Lt} \sim \left( \frac{\sigma - 1}{\sigma} \beta \right)^{\frac{1}{\beta+1}} \frac{H_{t}^{P}}{L} \varphi^{\mu(\psi-1)} \right)^{\frac{1}{\beta+1}} N_{t}^{\frac{\psi}{\beta+1}} \left( \varphi^{\mu-1} + (1 - G_{t}) w_{Lt}^{(\varepsilon-1)(1-\mu)} \right)^{\frac{1}{\beta+1}}.
\]
Following the assumption of Part B in Proposition 2, we assume that \( \lim (1 - G_{t}) N_{t}^{\frac{\psi}{\varepsilon(1-1-\mu)}} \exists \) exists and is finite. Suppose first that \( \limsup (1 - G_{t}) w_{Lt}^{(\varepsilon-1)(1-\mu)} = \infty \), then there must
exist a sequence of \( t \)'s, denoted \( t_n \) for which:

\[
\begin{align*}
\left(\left(\frac{\sigma - 1}{\sigma}\right)^{\frac{1}{1 - \beta}} \frac{H_P}{L} \varphi^{\mu - 1}\right)^{\frac{1}{1 + \beta(\sigma - 1)}} \left( (1 - G_{t_n}) N_{t_n}^{\psi}\right) \sim (w_{Lt_n})^\psi.
\end{align*}
\]

Yet, this implies

\[
(1 - G_{t_n}) w_{Lt_n}^{(\epsilon - 1)(1 - \mu)} \sim \left(\left(\frac{\sigma - 1}{\sigma}\right)^{\frac{1}{1 - \beta}} \frac{H_P}{L} \varphi^{\mu - 1}\right)^{(\epsilon - 1)(1 - \mu)} \left(1 - G_{t_n}\right)^{\frac{\psi(\epsilon - 1)(1 - \mu)}{1 + \beta(\sigma - 1)}}
\]

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

Consider now the possibility that \( \lim (1 - G_t) w_{Lt}^{(\epsilon - 1)(1 - \mu)} = 0 \), then (80) implies

\[
\begin{align*}
\left(\left(\frac{\sigma - 1}{\sigma}\right)^{\frac{1}{1 - \beta}} \frac{H_P}{L} \varphi^{\mu - 1}\right)^{\frac{1}{1 + \beta(\sigma - 1)}} \left(1 - G_{t_n}\right) N_{t_n}^{\psi} \sim \left(\left(\frac{\sigma - 1}{\sigma}\right)^{\frac{1}{1 - \beta}} \frac{H_P}{L} \varphi^{\mu - 1}\right)^{\frac{1}{1 + \beta(\sigma - 1)}} \left(1 - G_{t_m}\right) N_{t_m}^{\psi}.
\end{align*}
\]

Therefore we get that \( g^L = \frac{\psi}{\epsilon} g_N^{\psi} = \frac{1}{\epsilon} g_N^{Y} \).\(^{40}\)

Alternatively, \( \lim \sup (1 - G_t) w_{Lt}^{(\epsilon - 1)(1 - \mu)} \) is finite but strictly positive (given by \( \lambda_1 \)). In this case, there exists a sequence of \( t \)'s, denoted \( t_m \) such that

\[
\begin{align*}
\left(\left(\frac{\sigma - 1}{\sigma}\right)^{\frac{1}{1 - \beta}} \frac{H_P}{L} \varphi^{\mu - 1}\right)^{\frac{1}{1 + \beta(\sigma - 1)}} \left(1 - G_{t_m}\right) N_{t_m}^{\psi} \sim \left(\left(\frac{\sigma - 1}{\sigma}\right)^{\frac{1}{1 - \beta}} \frac{H_P}{L} \varphi^{\mu - 1}\right)^{\frac{1}{1 + \beta(\sigma - 1)}} \left(1 - G_{t_m}\right) N_{t_m}^{\psi}.
\end{align*}
\]

This leads to

\[
\begin{align*}
\lambda_1 \sim \left(\left(\frac{\sigma - 1}{\sigma}\right)^{\frac{1}{1 - \beta}} \frac{H_P}{L} \varphi^{\mu - 1}\right)^{\frac{1}{1 + \beta(\sigma - 1)}} \left(1 - G_{t_m}\right) N_{t_m}^{\psi(\epsilon - 1)(1 - \mu)}.
\end{align*}
\]

which is only possible if \( \lim (1 - G_t) N_{t}^{\psi(\epsilon - 1)(1 - \mu)} > 0 \). We denote such a limit by \( \lambda \).

\(^{40}\)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.
Then (80) leads to

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

which defines uniquely the limit of \( w_{Lt}^\epsilon, N_t^{-\psi} \). We then obtain that \( g_{\infty}^w = \frac{\psi}{\epsilon} g_{\infty}^N \). This completes the proof of part B).

**Subcase with \( \epsilon = \infty \).** Low skill wages are now defined as described in Appendix 7.1.2. First consider the case where \( G_\infty < 1 \), then Part A) immediately follows. Assume now that \( G_\infty = 1 \) and that \( \lim (1 - G_t) N_t^{\psi} \) exists and is finite. Note first that (75) implies that \( w_{Lt} \) must be bounded weakly above \( \tilde{\varphi} \) in the long-run. As a result, (76) leads to

\[ w_{Lt} \sim \left( \frac{\sigma - 1}{\sigma} \beta \tilde{\varphi}^{(1-\psi)} \right) \frac{1}{1-\beta} \frac{H_P^P}{L} \left( 1 - G_t \right) N_t^{\psi} \left( 1 + \varphi w_{Lt}^{\epsilon} \right) \frac{1}{1+\beta} \left( \sigma - 1 \right) \left( \sigma - 1 \right) \left( \varphi^{\mu} + 1 \right) \]

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 \( w_{L,\infty} = \tilde{\varphi} \).

**Case where \( G_\infty = 0 \) (Part C).** If \( \lim G_t = 0 \) then (72) implies that for \( \epsilon < \infty \):

\[ w_{Lt} \sim \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_P^P}{L} \right) \left( 1 - \beta \right) N_t^{\frac{1}{\sigma-1}} \left( G_t \left( 1 + \varphi w_{Lt}^{\epsilon} \right)^{\mu} + 1 \right) \left( 1 - \beta \right) \left( \varphi^{\mu} + 1 \right) \]

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:

\[ w_{Lt} \sim \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_P^P}{L} \right) \left( 1 - \beta \right) N_t^{\frac{1}{\sigma-1}} \left( G_t \varphi w_{Lt}^{\beta(\sigma-1)} + 1 \right) \left( 1 - \beta \right) \left( \varphi^{\mu} + 1 \right) \]

(82)

Note that if \( \epsilon = \infty \), then we must be in case iii) when \( G_\infty = 0 \) and (76) also directly implies (82) (as \( \varphi^{\mu} = \varphi^{\beta(\sigma-1)} \) in that case).

Assume that \( \lim_{t \to \infty} G_t N_t^{\beta} = \lambda \) exists and is finite. Then (82) implies:

\[ w_{Lt} N_t^{\frac{1}{\sigma-1}} \sim \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_P^P}{L} \right) \left( 1 - \beta \right) \left( \lambda \left( w_{Lt} N_t^{\frac{1}{\sigma-1}} \right)^{\beta(\sigma-1)} \varphi^{\mu} + 1 \right) \left( 1 - \beta \right) \left( \varphi^{\mu} + 1 \right) \]

69
which implies that \( \lim_{t \to \infty} w_t N_t^{\frac{1}{\sigma - 1}} \) exists and is finite as well. Therefore one gets that \( g^w_t \sim g^N_t / (\sigma - 1) \). Using (71) then immediately implies (17).

### 7.2.2 Sufficient conditions for Part A of Proposition 2

We prove the following Lemma:

**Lemma 1.** Consider processes \( [N_t]_{t=0}^{\infty}, [G_t]_{t=0}^{\infty} \) and \( [H^P_t]_{t=0}^{\infty} \), such that \( g^N_t \) and \( H^P_t \) 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 \( (G_t N_t) = \nu_{1,t} (1 - G_t) N_t + \nu_{2,t} \dot{N}_t \) such that \( \dot{G}_t = \nu_{1,t} (1 - G_t) - (G_t - \nu_{2,t}) g^N_t \). 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 \( \nu_{1,t} > \bar{\nu} \) 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 A combined HO-AR model

We now assume that product \( i \) is produced according to

\[
y(i) = ((b(i) l(i)) + \alpha(i) b(i) x(i))^{\beta} (b(i) h(i))^{1-\beta}
\]

instead of (2). That is we restrict attention for simplicity to the case where low-skill labor and machines are perfect substitute in automated firms and we normalize \( \bar{\varphi} \) to 1. We assume that \( b(i) = \exp(Bi) \) for some \( B > 0 \) and \( \zeta \in [0,1] \). \( \zeta \) represents the share of technological progress in new products which is TFP-augmenting while \( 1 - \alpha \) is the share which is purely labor augmenting (for both forms of labor). We assume that \( N_t \) grows exogenously and linearly, so that \( N_t = nt \) for some \( n > 0 \). Further, once invented a good has an exogenous probability \( \eta \) of becoming automated, and all high-skilled workers are hired in production \( (H^P = H) \).
We find that in this model low-skill and high-skill wages grow at the same rate if and only if the form of technological progress which is associated from moving from product $i$ to product $i'$ with a higher index is purely labor augmenting ($\varsigma = 0$). As soon as newer products also feature more productive machines if automated ($\varsigma > 0$) then low-skill wages will not grow at the same rate as output asymptotically. This result is in the spirit of Uzawa’s theorem but differs in so far as it refers to technological progress from one product to another instead of aggregate technological progress (which here features increasing the range of products and automating some of them). Formally, we establish:

**Proposition 8.** Low-skill and high-skill wages grow asymptotically at the same rate if and only if $\varsigma = 0$. Otherwise, high-skill wages grow asymptotically faster than low-skill wages: $g^{wH} > g^{wL}$.

**Proof.** We denote by $g(i, t)$ the share of products indexed by $i$ which are automated by time $t$. The Poisson process for automation implies that

$$g(i, t) = 1 - \exp\left(-\eta\left(t - \frac{i}{n}\right)\right) \quad \text{for} \quad t \geq \frac{i}{n}, \quad (83)$$

since a product $i$ must be born at time $t = i/n$ (and it is born non-automated).

The unit cost function of product $i$ can be written as:

$$c_i(w_L, w_H, \alpha(i)) = \left(\min\left(\frac{w_L}{b(i)^{1-\varsigma}}, \frac{1}{\alpha(i)}\right)\right)^{\beta} \frac{w_H^{1-\beta}}{b(i)^{1-\beta(1-\varsigma)}} \frac{1}{\beta^{\beta(1-\beta)1-\beta}}.$$

Therefore an automated firm ($\alpha(i) = 1$) will use machines instead of low-skill labor if $w_L > b(i)^{1-\varsigma}$. This implies that two cases must be considered, if $w_{Lt} > \exp(B(1-\varsigma)nt)$, then all automated firms will use machines. Otherwise, there exists a $I_t \in [0, N_t)$ such that $w_{Lt} = \exp(B(1-\varsigma)I_t)$ and automated firms with an index $i < I_t$ use machines while those with an index $i > I_t$ use low-skill workers instead. We fix $I_t = N_t$ if $w_{Lt} \geq \exp(B(1-\varsigma)nt$.

The resolution of the system then follows that in the baseline case, firms charge a mark-up $\sigma/(\sigma - 1)$ and revenues of firm $i$ are given by

$$R_i(w_L, w_H, \alpha(i)) = \left(\frac{\sigma - 1}{\sigma}\right)^{\sigma-1} c_i(w_L, w_H, \alpha(i))^{1-\sigma} Y.$$

A share $(1-\beta)(\sigma - 1)/\sigma$ of firms’ revenues accrue to high-skill workers, while low-skill workers obtain a share $\beta(\sigma - 1)/\sigma$ if the firm is non-automated or has an index $i > I_t$.
and 0 otherwise. Therefore

\[ w_{Ht} = (1 - \beta) \frac{\sigma - 1}{\sigma} \int_0^{N_t} (1 - g(i, t)) R_i(w_{Lt}, w_{Ht}, 0) + g(i, t) R_i(w_{Lt}, w_{Ht}, 1) \, di, \]

\[ w_{Lt} = \beta \frac{\sigma - 1}{\sigma} \int_0^{N_t} (1 - g(i, t)) R_i(w_L, w_H, 0) + 1_{i > I_t} g(i, t) R_i(w_{Lt}, w_{Ht}, 1) \, di, \]

where \( 1_{i > I_t} \) denotes the index function for \( i > I_t \). Taking the ratio between these two expressions, we obtain:

\[ \frac{w_{Ht}}{w_{Lt}} = 1 - \beta \frac{\sigma - 1}{\sigma} \left[ \int_0^{N_t} 1_{i < I_t} g(i, t) b(i) (1 - \beta(1 - \varsigma)) (\sigma - 1) \, di \right], \]

which traces the relative demand curve (and replaces (10)).

Similarly, using the price normalization, we obtain:

\[ \frac{\sigma}{\sigma - 1} \frac{w_{Ht}^{1-\beta}}{w_{Lt}^{1-\beta}} \left( \int_0^{N_t} \left( (1 - 1_{i < I_t} g(i, t)) w_{Lt}^{(1-\sigma)} b(i) (1 - \beta(1 - \varsigma)) (\sigma - 1) \right) \, di \right)^{\frac{1}{\sigma} - 1} = 1, \]

which replaces (11). We must then consider two cases in turn: \( I_t = N_t \) and \( I_t < N_t \).

**Case 1: \( I_t = N_t \).** Using (83) and the expression for \( b(i) \) we can compute the integral in (84) and obtain:

\[ \frac{w_{Ht}}{w_{Lt}} = 1 - \beta \frac{\sigma - 1}{\sigma} \left[ 1 + w_{Lt}^{\beta(\sigma-1)} \left( \frac{\exp(B(1-\beta(1-\varsigma))(\sigma-1)nt-1) - \exp(B(1-\beta(1-\varsigma))(\sigma-1)nt-\exp(-\eta t))}{\exp(B(\sigma-1)nt) - \exp(-\eta t) \left( B(\sigma-1) + \frac{\eta}{n} \right)} \right) \right]. \]

For \( t \) large enough (and since \( (1 - \beta(1 - \varsigma)) (\sigma - 1) > 0 \), we get:

\[ \frac{w_{Ht}}{w_{Lt}} \sim 1 - \beta \left[ 1 + \frac{\left( B (\sigma - 1) + \frac{\eta}{n} \right) \frac{n}{\sigma} (w_L \exp(-B (1 - \varsigma) nt))^{\beta(\sigma-1)}}{\left( B (1 - \beta (1 - \varsigma)) (\sigma - 1) \right) \left( B (1 - \beta (1 - \varsigma)) (\sigma - 1) + \frac{\eta}{n} \right)} \right]. \]
Similarly (85) gives

$$\frac{\sigma}{\sigma - 1} \beta^2 (1 - \beta)^{1-\beta} = \left( w_{Lt}^{\beta(1-\sigma)} \frac{\exp(B(\sigma-1)nt) - \exp(-nt)}{B(\sigma-1) + \frac{n}{\sigma}} + \frac{\exp(B(1-\beta(1-\varsigma)))(\sigma-1)nt-1}{B(1-\beta(1-\varsigma))(\sigma-1)(\sigma-1)+\frac{n}{\sigma}} \right) \frac{1}{1-\beta}. $$

From this, we obtain:

$$w_{Ht} \sim \exp \left( B \frac{(1-\beta)(1-\varsigma)}{1-\beta} nt \right) \left(1 - \beta \right) \left( \frac{\sigma - 1}{\sigma} \beta^2 \right)^{1-\sigma} $$

(87)

with \( \psi \equiv 1 / [(\sigma - 1)(1 - \beta)] \) as before.

Since \( I_t = N_t \), then we must have \( w_{Lt} \geq \exp(B(1-\varsigma)nt) \). Therefore (87) implies that

$$w_{Ht} = O \left( \exp \left( B \frac{(1-\beta)(1-\varsigma)}{1-\beta} nt \right) \right).$$

Further (86) implies that

$$\frac{w_{Ht}}{w_{Lt}} = O \left( w_L \exp(-B(1-\varsigma)nt)^{\beta(\sigma-1)} \right).$$

from which we get that

$$w_{Lt} = O \left( \exp \left( \left( \frac{1-\beta(1-\varsigma) + (1-\beta) \beta (\sigma-1)(1-\varsigma)}{1+\beta(\sigma-1)} (1 - \beta) \right) Bnt \right) \right).$$

We need to verify that \( w_{Lt} \geq \exp(B(1-\varsigma)nt) \). Note that

$$\left( \frac{1-\beta(1-\varsigma) + (1-\beta) \beta (\sigma-1)(1-\varsigma)}{1+\beta(\sigma-1)} (1 - \beta) \right) Bn \geq B(1-\varsigma)n \iff \varsigma \geq 0.$$

Therefore, if \( \varsigma > 0 \), then \( w_{Lt} > \exp(B(1-\varsigma)nt) \) is verified for sure for large \( t \), and we get that \( g_{w_H}^w \) and \( g_{w_L}^w \) exist with \( g_{w_H}^w = \frac{(1-\beta(1-\varsigma))Bn}{1-\beta} \) and \( g_{w_L}^w = \frac{1-\beta(1-\varsigma) + (1-\beta) \beta (\sigma-1)(1-\varsigma)}{(1+\beta(\sigma-1))(1-\beta)} Bn \).
and we can verify that \( g_{\infty}^w > g_{\infty}^L \).

In contrast if \( \zeta = 0 \), then we have \( g_{\infty}^w = g_{\infty}^L = Bn \) but this case only applies if there exist constant \( \overline{w}_H \) and \( \overline{w}_L \) such that

\[
\overline{w}_H = (1 - \beta) \left( \frac{\sigma - 1}{\sigma} \beta^\gamma \right)^{1/\beta} \left( \frac{\overline{w}_L^{\beta(1-\sigma)}}{B(\sigma - 1) + \frac{\eta}{n}} + \frac{\eta}{n} \right)^\psi,
\]

\[
\frac{\overline{w}_H H}{\overline{w}_L L} = \frac{1 - \beta}{\beta} \left[ 1 + \overline{w}_L^{\beta(\sigma - 1)} \frac{(B(\sigma - 1) + \frac{\eta}{n}) \frac{\eta}{n}}{B(1 - \beta)(\sigma - 1) (B(1 - \beta)(\sigma - 1) + \frac{\eta}{n})} \right],
\]

with \( \overline{w}_L \geq 1 \), in which case \( w_{Ht} \sim \overline{w}_H \exp (Bnt) \) and \( w_{Lt} \sim \overline{w}_L \exp (Bnt) \) and \( I_t = N_t \) is a possibility.

**Case 2**: \( I_t < N_t \). Then (84) implies

\[
\frac{w_{Ht} H}{w_{Lt} L} \sim \frac{1 - \beta}{\beta} \left[ 1 + w_{Lt}^{\beta(\sigma - 1)} \frac{\exp(B(1 - \beta(1 - \zeta))(\sigma - 1)I_t - 1)}{B(1 - \beta(1 - \zeta))(\sigma - 1)} - \exp \left( - \frac{\eta}{n} nt \right) \frac{\exp \left( B(1 - \beta(1 - \zeta))(\sigma - 1) + \frac{\eta}{n} \right) I_t - 1}{B(1 - \beta(1 - \zeta))(\sigma - 1) + \frac{\eta}{n}} \right] \exp \left( - \frac{\eta}{n} nt \right). \]

Note that we must have \( \lim I_t = \infty \), otherwise, there would be periods where \( w_{Lt} \) remain bounded even though an arbitrarily large number of products use low-skill workers.

Therefore, using \( w_{Lt} = \exp \left( (1 - \zeta) BI_t \right) \), the previous equation leads to:

\[
\frac{w_{Ht} H}{w_{Lt} L} \sim \frac{1 - \beta}{\beta} \left[ 1 + \frac{1}{B(1 - \beta)(\sigma - 1)} - \exp \left( - \frac{\eta}{n} nt \right) \frac{1}{B(1 - \beta(1 - \zeta))(\sigma - 1) + \frac{\eta}{n}} \right],
\]

Since \( nt \geq I_t \), then we must have that \( w_{Ht} = O \left( w_{Lt} \right) = O \left( \exp \left( (1 - \zeta) BI_t \right) \right) \). Similarly (85) now implies:

\[
\frac{\sigma w_{Ht}^{\beta - 1}}{(\sigma - 1) \beta \beta^{1 - \beta}} \exp \left( (1 - (1 - \zeta) \beta) BI_t \right) \left( \frac{\exp \left( B(\sigma - 1)(nt - I_t) - 1 \right)}{B(\sigma - 1)} + \frac{\exp \left( - \frac{\eta}{n} (nt - I_t) \right) \exp \left( B(\sigma - 1) + \frac{\eta}{n} \right) (nt - I_t)}{B(1 - \beta) \beta \beta^{1 - \beta} \left( \exp \left( B(\sigma - 1)(nt - I_t) - 1 \right) + \frac{\exp \left( - \frac{\eta}{n} (nt - I_t) \right) \exp \left( B(\sigma - 1) + \frac{\eta}{n} \right) (nt - I_t)}{B(1 - \beta) \beta \beta^{1 - \beta}} \right)} \right) \frac{1}{\beta^{1 - \beta}}.
\]

Using that \( I_t \to \infty \) and that \( w_{Lt} = \exp \left( (1 - \zeta) BI_t \right) \), we then get

\[
\frac{\sigma w_{Ht}^{\beta - 1}}{(\sigma - 1) \beta \beta^{1 - \beta}} \sim \exp \left( (1 - (1 - \zeta) \beta) BI_t \right) \left( \frac{\exp \left( B(\sigma - 1)(nt - I_t) - 1 \right)}{B(\sigma - 1)} + \frac{\exp \left( - \frac{\eta}{n} (nt - I_t) \right) \exp \left( B(\sigma - 1) + \frac{\eta}{n} \right) (nt - I_t)}{B(1 - \beta) \beta \beta^{1 - \beta} \left( \exp \left( B(\sigma - 1)(nt - I_t) - 1 \right) + \frac{\exp \left( - \frac{\eta}{n} (nt - I_t) \right) \exp \left( B(\sigma - 1) + \frac{\eta}{n} \right) (nt - I_t)}{B(1 - \beta) \beta \beta^{1 - \beta}} \right)} \right) \frac{1}{\beta^{1 - \beta}}.
\]
The left-hand side is of order \( \exp((1 - \zeta - \beta (1 - \zeta)) BI_t) \) while the the right-hand side is of order at least \( \exp((1 - (1 - \zeta) \beta) BI_t) \), therefore this situation is only possible if \( \zeta = 0 \). Moreover, if \( nt - I_t \) is unbounded then the ratio of the right-hand side to the left-hand side is also unbounded, which is a contradiction. Therefore it must be that \( nt - I_t \) remains bounded. In that case, we must then have that \( w_{Lt} \) and \( w_{Ht} \) are of the same order as \( \exp(B I_t) \).

We then obtain an equilibrium with \( I_t < N_t \) if there exist \( \bar{w}_H, \bar{w}_L < 1 \) (which implies \( \lim nt - I_t = -\log \bar{w}_L \)) such that:

\[
\frac{\bar{w}_H H}{\bar{w}_L L} = \frac{1 - \beta}{\beta} \left[ 1 + \frac{1}{B(1 - \beta)(\sigma - 1)} - \frac{\psi(w_L^\eta)}{B(1 - \beta)(\sigma - 1) + \frac{\eta}{n}} \right],
\]

\[\tag{90}\]

\[\bar{w}_H = (1 - \beta) \left( \frac{\sigma - 1}{\sigma} \beta^\alpha \right) \frac{1}{\bar{w}_L} \left( \frac{\psi(w_L^\eta)}{B(\sigma - 1)} - \frac{\psi(w_H^\eta B \beta(\sigma - 1))}{B(\sigma - 1) + \frac{\eta}{n}} \right)^\psi.
\]

\[\tag{91}\]

(90) can be rewritten as \( \bar{w}_H = \frac{1 - \beta}{\beta} L \bar{w}_L s_1(\bar{w}_L) \) with

\[s_1(\bar{w}_L) \equiv \bar{w}_L \left[ 1 + \frac{1}{B(1 - \beta)(\sigma - 1)} - \frac{\psi(w_L^\eta)}{B(1 - \beta)(\sigma - 1) + \frac{\eta}{n}} \right].
\]

We obtain (after a lot of algebra):

\[s'_1(\bar{w}_L) = \left\{ \begin{array}{l}
\left[ \frac{n}{\eta} \frac{\frac{n}{\eta} + 1}{(B(\sigma - 1) + \frac{\eta}{n}) (B(1 - \beta)(\sigma - 1) + \frac{\eta}{n})} + \frac{1}{B(1 - \beta)(\sigma - 1)} \right] + \frac{\beta(\sigma - 1)}{B(\sigma - 1) + \frac{\eta}{n}} \frac{\psi(w_L^\eta B \beta(\sigma - 1))}{B(\sigma - 1) + \frac{\eta}{n}} \\
+ \left[ \frac{\psi(w_L^\eta B \beta(\sigma - 1))}{B(\sigma - 1) + \frac{\eta}{n}} \right] \left[ \frac{\psi(w_L^\eta B \beta(\sigma - 1))}{B(\sigma - 1) + \frac{\eta}{n}} \right] \left[ \frac{\psi(w_L^\eta B \beta(\sigma - 1))}{B(\sigma - 1) + \frac{\eta}{n}} \right] \left[ \frac{\psi(w_L^\eta B \beta(\sigma - 1))}{B(\sigma - 1) + \frac{\eta}{n}} \right]
\end{array} \right.,
\]

which is positive when \( \bar{w}_L \leq 1 \) (as then \( \psi(w_L^\eta) \leq 1 \geq 0 \) and \( 1 - \psi(w_L^\eta) \geq 0 \)), so that \( s_1 \) is increasing in \( \bar{w}_L \) for \( \bar{w}_L \leq 1 \). Similarly (91) can be rewritten as \( \bar{w}_H = (1 - \beta) \left( \frac{\sigma - 1}{\sigma} \beta^\alpha \right)^ {1 - \beta} s_2(\bar{w}_L)^\psi, \)
with
\[ s_2 (\bar{w}_L) \equiv \bar{w}_L^{(\sigma-1)(1-\beta)} \left( \frac{\bar{w}_L^{(1-\sigma)} - 1}{B (\sigma - 1)} - \frac{\bar{w}_L^n B \beta (\sigma - 1)}{B (\sigma - 1) + \frac{n}{\sigma}} \right) + \frac{1}{B (1 - \beta) (\sigma - 1)}. \]

We get
\[ s'_2 (\bar{w}_L) = -\bar{w}_L^{(\sigma-1)(1-\beta)-1} \left( \frac{B (\sigma - 1) \beta}{B (\sigma - 1) + \frac{n}{\sigma}} \right) \left( \frac{\bar{w}_L^n B}{B} + \beta \frac{\bar{w}_L^{1-\sigma} - 1}{B} \right), \]

which is negative for \( \bar{w}_L \leq 1 \). Therefore (89) and (90) together trace a positive relationship between \( \bar{w}_H \) and \( \bar{w}_L \) (it is easy to show continuity between the two expressions) and similarly (88) and (91) trace a positive relationship (continuity is also easily verified), checking the limits when \( \bar{w}_L \to 0, \infty \), we obtain that there exist a single solution. This ensures that asymptotically when \( \varsigma = 0 \), the asymptotic equilibrium described above exists, and depending on other parameters it is in case 1 or case 2.

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, \( \sigma \). 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:

\[ y_k(i) = \left[ l_k(i) \beta_k + \alpha(i) (\tilde{\phi x}_k(i)) \beta_k \right] h_k(i) h_1(i)^{1-\beta_k}, \]

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 \( p_x \), and allowing for the natural extensions of market clearing conditions, we derive below the following generalization of Proposition 2 (where \( \psi_k = (\sigma - 1)(1 - \beta_k)^{-1} \)).

**Proposition 9.** Consider three processes \( [N_t]_{t=0}^\infty, [G_t]_{t=0}^\infty \) and \( [H_t^p]_{t=0}^\infty \) where \( (N_t, G_t, H_t^p) \in (0, \infty) \times [0, 1] \times (0, \infty) \) for all \( t \). Assume that \( G_t, g^N_t \) and \( H_t^p \) all admit strictly positive limits, then:

\[
\begin{align*}
g_\infty^p &= -\psi_2 (\beta_2 - \beta_1) g^N_\infty \\
g_\infty^{GDP} &= \left[ \psi_1 + \psi_1 \frac{\beta_1 (\beta_2 - \beta_1)}{1 - \beta_2} \right] g^N_\infty,
\end{align*}
\]

and if \( G_\infty < 1 \) then the asymptotic growth rate of \( w_L \) is \( w_\infty^L = \frac{1}{1 + \psi_1 (\sigma - 1)} \frac{1 - \beta_2 + \beta_1 (\beta_2 - \beta_1)(1 - \psi_1^{-1})}{1 - \beta_2 + \beta_1 (\beta_2 - \beta_1)} g_\infty^{GDP}. \)

Proposition 9 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 \( g^N_\infty \), 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 for higher values of the elasticity of substitution between products, \( \sigma \), 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 (4)

\[ g_\infty^p = \frac{1}{1 + \psi_1 (\sigma - 1)} \frac{1 - \beta_2 + \beta_1 (\beta_2 - \beta_1)(1 - \psi_1^{-1})}{1 - \beta_2 + \beta_1 (\beta_2 - \beta_1)} g_\infty^{GDP}. \]

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

\[ c_k(\alpha(i)) = \beta_k \alpha(1 - \beta_k)^{-1}(1 - \beta_k) (w_L^{1-\epsilon} + \varphi(p^x)^{1-\epsilon} \alpha(i))^{\frac{\beta_k}{1-\epsilon}} w_H^{1-\beta_k}, \] (94)

for \( k \in \{1, 2\} \) indexing, respectively, the production of final good and machines. As before aggregating (94) and the price normalization gives a “productivity” condition, which replaces (11).

\[ \left( G \left( w_L^{1-\epsilon} + \varphi(p^x)^{1-\epsilon}\right)^{\mu_1} + (1 - G) w_L^{\beta_1(1-\sigma)} \right)^{\frac{1}{1-\sigma}} w_H^{1-\beta_1} = \frac{\sigma - 1}{\sigma} \beta_1 (1 - \beta_1)^{1-\beta_1} N^{\frac{1}{\sigma-1}}, \] (95)

where we generalize the definition of \( \mu_1 \): \( \mu_k \equiv \frac{\beta_k(\sigma-1)}{\epsilon-1} \). Following the same methodology for the production of machines, we get

\[ \left( G \left( w_L^{1-\epsilon} + \varphi(p^x)^{1-\epsilon}\right)^{\mu_2} + (1 - G) w_L^{\beta_2(1-\sigma)} \right)^{\frac{1}{1-\sigma}} w_H^{1-\beta_2} = \frac{\sigma - 1}{\sigma} \beta_2 (1 - \beta_2)^{1-\beta_2} N^{\frac{1}{\sigma-1}} p^x. \] (96)

Taking the ratio between these two expressions, we get

\[ \frac{\left( G \left( \frac{w_L}{p^x} \right)^{1-\epsilon} + \varphi \right)^{\mu_2} + (1 - G) \left( \frac{w_L}{p^x} \right)^{\beta_2(1-\sigma)} \right)^{\frac{1}{1-\sigma}} w_H^{\beta_2-1-\beta_2}}{\left( G \left( \frac{w_L}{p^x} \right)^{1-\epsilon} + \varphi \right)^{\mu_1} + (1 - G) \left( \frac{w_L}{p^x} \right)^{\beta_1(1-\sigma)} \right)^{\frac{1}{1-\sigma}} w_H^{\beta_1-1-\beta_1}} = \frac{\beta_2 (1 - \beta_2)^{1-\beta_2} (p^x)^{1-\beta_2+\beta_1}}{\beta_1 (1 - \beta_1)^{1-\beta_1}}. \] (97)

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

\[ \nu_{k,x}(\alpha(i)) = \frac{\sigma - 1}{\sigma} \alpha(i) \beta_k \varphi(p^x)^{1-\epsilon} w_L^{1-\epsilon} + \varphi(p^x)^{1-\epsilon}, \] (98)

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

\[ p^x X = N G \left( R_1(1) \nu_{1,x}(1) + R_2(1) \nu_{2,x}(1) \right). \] (99)

The zero profit condition in the machines sector gives

\[ p^x X = N \left( GR_2(1) + (1 - G) R_2(0) \right). \] (100)
Revenues themselves are given by

\[ R_1 (\alpha (i)) = \left( \frac{\sigma - 1}{\sigma} \right)^{\sigma - 1} c_1 (\alpha (i))^{1 - \sigma} Y \]
and

\[ R_2 (\alpha (i)) = \left( \frac{\sigma - 1}{\sigma} \right)^{\sigma - 1} c_2 (\alpha (i))^{1 - \sigma} p^x X, \]

so that (7) still holds but separately for revenues occurring from each activity and with \( \mu_k \) replacing \( \mu \). Combining (7), (98), (99) and (100), we get

\[ \left( G \left( 1 - \frac{\sigma - 1}{\sigma} \beta_2 \frac{\varphi (p^x)^{1 - \epsilon}}{w_L (1 + \varphi (p^x)^{1 - \epsilon}) - \epsilon - 1} \right) + (1 - G) \left( 1 + \varphi \left( \frac{w_L}{p^x} \right)^{\epsilon - 1} \right)^{-\mu_2} \right) \frac{R_2(1)}{R_1(1)}, \]

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:

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

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

\[ w_L L = N (GR_1(1)\nu_{1,l}(1) + (1 - G)R_1(0)\nu_{1,l}(0) + GR_2(1)\nu_{2,l}(1) + (1 - G)R_2(0)\nu_{2,l}(0)). \]

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

\[ w_H H^P = N (\nu_{1,h} (GR_1(1) + (1 - G)R_1(0)) + \nu_{2,h} (GR_2(1) + (1 - G)R_2(0))), \]
Take the ratio between \((103)\) and \((104)\), and use \((7)\) to obtain:

\[
\frac{w_L L}{w_H H^P} = \frac{\beta_1 \left( G \left( 1 + \varphi \left( \frac{w_L}{p^z} \right)^{\epsilon - 1} \right)^{-1} + (1 - G) \left( 1 + \varphi \left( \frac{w_L}{p^z} \right)^{\epsilon - 1} \right)^{-\mu_1} \right)}{+ \beta_2 R_2(1) \left( G \left( 1 + \varphi \left( \frac{w_L}{p^z} \right)^{\epsilon - 1} \right)^{-1} + (1 - G) \left( 1 + \varphi \left( \frac{w_L}{p^z} \right)^{\epsilon - 1} \right)^{-\mu_2} \right)}.
\]

Together \((95)\), \((97)\), \((102)\) and \((105)\) determine \(w_L, w_H, p^z\) 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 \(w_{L\infty} / p^z_{\infty} = \infty\) even if \(w_{L\infty} < \infty\). Applying this to \((97)\), we get

\[
(p_t^z)^{1-\beta_2 + \beta_1} \sim \frac{\beta_1^{1-\beta_1} \beta_2^{1-\beta_2}}{\beta_1^{1-\beta_1} \beta_2^{1-\beta_2}} \frac{\varphi^{\epsilon - 1}}{w_H^{1-\beta_2}}.
\]

Further plugging this last relationship in \((95)\), we get:

\[
w_{Ht} \sim \left( \frac{\sigma - 1}{\sigma} \right)^{1+\beta_1} \varphi^{\psi_2 \mu_1} \beta_1^{1-\beta_1} \beta_2^{1-\beta_2} \left( \frac{\beta_1}{\beta_2} \right) \left( \frac{\beta_2}{\beta_1} \right)
\]

Hence

\[
g_{w_H} = \psi_1 \left( 1 + \frac{\beta_1 (\beta_2 - \beta_1)}{(1 - \beta_2)} \right) g_{N}^{\infty}.
\]

Through \((102)\), the revenues of the machines sector and the final good sector are of the same order, which implies that \(Y, p^z X\) and \(w_H\) grow at the same rate. Therefore

\[
g_{GDP} = g_{Y} = g_{w_H} = \psi_1 \left( 1 + \frac{\beta_1 (\beta_2 - \beta_1)}{(1 - \beta_2)} \right) g_{N}^{\infty}.
\]

In fact \((102)\) gives

\[
\frac{R_{2,t}(1)}{R_{1,t}(1)} \sim \frac{\frac{\sigma - 1}{\sigma} \beta_1}{1 - \frac{\sigma - 1}{\sigma} \beta_2}.
\]
Using (106) and (107), one further gets:

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

therefore

\[ g_p^\infty = -\psi_2 (\beta_2 - \beta_1) g_\infty^N < 0, \quad (110) \]

since \( \beta_2 > \beta_1 \). Using that \( w_{L\infty}/p_t^x = \infty \) and (109) in (105) leads to:

\[ w_{Lt} \left( \frac{w_{Lt}}{p_t^x} \right)^{\epsilon - 1} \sim \frac{w_{Lt} H_t^P \left( \beta_1 \left( G_t + (1 - G_t) \left( \varphi \left( \frac{w_{Lt}}{p_t^x} \right)^{\epsilon - 1} \right)^{1 - \mu_1} \right) \right) + \beta_2 \frac{\sigma - 1}{\sigma - 1 - \beta_2} \left( G_t + (1 - G_t) \left( \varphi \left( \frac{w_{Lt}}{p_t^x} \right)^{\epsilon - 1} \right)^{1 - \mu_2} \right)}{\varphi G_t L \left( 1 - \beta_1 + (1 - \beta_2) \frac{\sigma - 1}{\sigma - 1 - \beta_2} \right)}. \quad (111) \]

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 \( G_\infty \) (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 7.2. If \( G_\infty < 1 \), then (111) implies

\[ w_{Lt}^{1 + \beta_1 (\sigma - 1)} \sim \left( p_t^x (\sigma - 1) \beta_1 \right) \frac{w_{Lt} H_t^P \beta_1 (1 - G_t)}{\varphi^\mu G_t L \left( 1 - \beta_1 + (1 - \beta_2) \frac{\sigma - 1}{\sigma - 1 - \beta_2} \right)}, \quad (112) \]

which, together with (108) and (110) gives (93).

Alternatively assume that \( G_\infty = 1 \) and that \( \lim (1 - G_t) N_t^{\psi_2 (1 - \mu_1) \frac{\epsilon - 1}{\epsilon - 1}} \) exists and is finite. Suppose first that \( \lim sup (1 - G_t) \left( \frac{w_{Lt}}{p_t^x} \right)^{(\epsilon - 1)(1 - \mu_1)} = \infty \), then there must be a sub-sequence where (112) is satisfied, which with (108) and (110) leads to a contradiction with the assumption that \( \lim (1 - G_t) N_t^{\psi_2 (1 - \mu_1) \frac{\epsilon - 1}{\epsilon - 1}} \) exists and is finite.

If \( \lim (1 - G) \left( \frac{w_{Lt}}{p_t^x} \right)^{(\epsilon - 1)(1 - \mu_1)} = 0 \), then (111) gives

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

81
which implies with (108) and (110) that:

\[ g^*_\infty = \frac{1}{\epsilon} \left( 1 - \frac{(\beta_2 - \beta_1)(\epsilon - 1)}{(1 - \beta_2 + \beta_1)} \right) g^*_\text{GDP}. \]  

(113)

Finally, if \( \limsup (1 - G_t) w^\epsilon_t(1-\mu) \) is finite but strictly positive, then as in Appendix 7.2.1, one can show that this requires that \( \lim (1 - G_t) N^\epsilon_t(1-\mu) > 0 \), from which we can derive that (113) also holds in that case. This proves Proposition 9 and the associated footnote in the imperfect substitutes case.

**Perfect substitutes case.** In the perfect substitutes case, (95) becomes:

\[
\left( G\tilde{\varphi}^{\beta_1(\sigma-1)}(p^x)^{\beta_1(1-\sigma)} + (1 - G)w^\beta_L(1-\sigma) \right)^{\frac{1}{1-\sigma}} w^1_L^{1-\beta_1} = \frac{\sigma - 1}{\sigma} \beta_1^1 (1 - \beta_1)^{1 - \beta_1} N^{\frac{1}{\sigma - 1}} \text{ for } w_L > p^x/\tilde{\varphi},
\]

(114)

\[
w^1_L w_H^{1-\beta_1} = \frac{\sigma - 1}{\sigma} \beta_1^1 (1 - \beta_1)^{1 - \beta_1} N^{\frac{1}{\sigma - 1}} \text{ for } w_L < p^x;
\]

(115)

(97) becomes

\[
\left( G\tilde{\varphi}^{\beta_2(\sigma-1)}(p^x)^{\beta_2(1-\sigma)} + (1 - G)w^\beta_L(1-\sigma) \right)^{\frac{1}{1-\sigma}} w^1_H^{1-\beta_2} = \frac{\beta_2^1 (1 - \beta_2)^{1 - \beta_2} p^x}{\beta_1^1 (1 - \beta_1)^{1 - \beta_1}} \text{ for } w_L > p^x/\tilde{\varphi},
\]

(116)

\[
p_x = \frac{\beta_1^1 (1 - \beta_1)^{1 - \beta_1} w^1_L w^1_H^{1-\beta_1}}{\beta_2^1 (1 - \beta_2)^{1 - \beta_2} w^1_L w^1_H^{1-\beta_2}} \text{ for } w_L < p^x/\tilde{\varphi};
\]

(117)

(102) becomes

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

(118)

with \( R_2(1) = 0 \) for \( w_L < p^x/\tilde{\varphi} \); and (105) becomes

\[
\frac{w^*_L}{w^*_H p^x} = (1 - G) \left\{ \beta_1 \left( \tilde{\varphi}^{\frac{w^*_L}{p^x}} \right)^{\beta_1(1-\sigma)} + \beta_2 \frac{R_2(1)}{R_1(1)} \left( \tilde{\varphi}^{\frac{w^*_L}{p^x}} \right)^{\beta_2(1-\sigma)} \right\} \text{ for } w_L > p^x/\tilde{\varphi},
\]

(119)
\[
\frac{w_L}{w_H H^P} = \frac{\beta_1}{1 - \beta_1} \quad \text{for } w_L < \frac{p^*}{\bar{\varphi}}. \tag{120}
\]
Together (115), (117) and (120) show that we must have \( w_{Lt} \geq \frac{p^*}{\bar{\varphi}} \) for \( t \) large enough, which delivers (108) and (110).

Assume that \( G_\infty < 1 \), then (119) gives (112) from which we get that (93) is satisfied.

Now consider the case where \( G_\infty = 1 \) and \( \lim (1 - G_t) N_t^{\psi_2} \) exists and is finite. Then (119) and (118) imply

\[
w_{Lt} \sim (1 - G_t) w_{Ht} \left( \frac{\bar{\varphi} w_{Lt}}{p^*_t} \right)^{\beta_1(1-\sigma)} \left( \beta_1 + \beta_2 \frac{\bar{\varphi} w_{Lt}}{p^*_t} \left( \frac{w_{Lt}}{p^*_t} \right)^{-(\beta_2 - \beta_1)(\sigma - 1)} \right) \frac{H^P_t}{L} \text{ for } w_L > \frac{p^*}{\bar{\varphi}}.
\]

We can then derive that \( \frac{\bar{\varphi} w_{Lt}}{p^*_t} \) must have a finite (and positive) limit, so that

\[
g_{\infty} = g^x_{\infty} = -\frac{\beta_2 - \beta_1}{1 - \beta_2 + \beta_1} g^{GDP}_{\infty}.
\]

This proves Proposition 9 and its associated footnote in the perfect substitutes case. \( \square \)

### 7.5 Proofs and analytical results for the baseline dynamic model

#### 7.5.1 Uniqueness of the steady state

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

\[
f \left( g^{N*} \right) = \frac{1 - \kappa}{\kappa} G^* \widehat{h}^{A*} \left( \frac{1}{\kappa} \left( \frac{\kappa}{\gamma(1-\kappa)} \right)^{1-\kappa} \left( \frac{\kappa}{\gamma(1-\kappa)} \right)^{\frac{1}{\gamma}} + \frac{1}{\gamma} \right), \tag{121}
\]

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 \( \bar{\kappa} = 0 \), using (59), (58), we get:

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

83
Therefore
\[
\frac{d}{dg^N} \left( G^* \tilde{h}^A \right) = \frac{\eta (\gamma (1-\gamma))^{\kappa+1} (\rho + ((\theta - 1) \psi + 1) g^N)^{\kappa}}{(\eta (\gamma (1-\gamma))^{\gamma (\rho + ((\theta - 1) \psi + 1) g^N)^{\kappa} + dN^{2})} \cdot (\theta - 1) \psi + 1) (\rho + ((\theta - 1) \psi + 1) g^N)^{\kappa})
\]

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

### 7.5.2 Transitional dynamics and the first phase

Here we prove Proposition 4A. Combining (29) and (28), we can write:

\[
N_t h_t^A = \left( \kappa \eta \tilde{G}_t^\kappa \left( \int_t^\infty \exp \left( - \int_t^\tau r_u du \right) \left( \frac{N_t}{w_{Ht}} \left( \pi_t^A - \pi_t^N \right) d\tau - \frac{1 - \kappa}{\kappa} \frac{N_t}{w_{Ht}} \left( \frac{\pi_t^A - \pi_t^N}{\pi_t^A - \pi_t^N} \right) d\tau \right) \right) \right)^{1-\kappa}.
\]

Using (8) 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:

\[
\tilde{h}_t^A = \left( \kappa \eta \tilde{G}_t^\kappa \left( \int_t^\infty \exp \left( - \int_t^\tau r_u du \right) \left( \psi H_t^P \left( \pi_t^A - \pi_t^N \right) d\tau - \frac{1 - \kappa}{\kappa} \frac{N_t}{w_{Ht}} \left( \frac{\pi_t^A - \pi_t^N}{\pi_t^A - \pi_t^N} \right) d\tau \right) \right) \right)^{1-\kappa}.
\]

Recalling (7), we can write:

\[
\tilde{h}_t^A = \left( \kappa \eta \tilde{G}_t^\kappa \left( \int_t^\infty \left( \psi H_t^P \left( \frac{(1 + \varphi w_{Lt}^{\epsilon-1})^{\epsilon-1}}{G_t(1 + \varphi w_{Lt}^{\epsilon-1})^{\epsilon+1}} \exp \left( \int_t^\tau \left( g_u^N - r_u \right) du \right) \right) \right) \right)^{1-\kappa}.
\]

Consider a fixed \( \tilde{t} > 0 \). Then for an arbitrarily large \( T \), if \( w_{L0} \) is sufficiently small relative to \( \tilde{\varphi}^{-1} \), we will have that \( w_{Lt} \) is small relative to \( \tilde{\varphi}^{-1} \) over \( (0, \tilde{t} + T) \). For any \( \tau \in (0, \tilde{t} + T) \), we have that \( \tilde{G}_t(1 + \varphi w_{Lt}^{\epsilon-1})^{\epsilon+1} = \mu \varphi w_{L\tau}^{\epsilon-1} + o \left( \varphi w_{L\tau}^{\epsilon-1} \right) \). The notation \( O(z) \) denotes negligible relative to \( z \) (that is \( f(z) = o(z) \), if \( \lim f(z)/z = 0 \) and \( O(z) \) will denote of the same order or negligible in front of \( z \) \( f(z) = O(z) \) if \( \lim sup |f(z)/z| < \))
Then for any $t \in (0, \hat{t})$

$$\left(\hat{h}_t^A\right)^{1-\kappa} \leq \kappa \eta G_t^{\tilde{\kappa}} \left( \int_t^{\tilde{t}+T} \psi H_t^P \left( \mu \varphi w_{Lt}^{t-1} + o \left( \varphi w_{Lt}^{t-1} \right) \right) \exp \left( \int_t^\tau \left( g_u^{\pi N} - r_u \right) du \right) d\tau \right)$$

$$+ \int_{t+T}^{\infty} \psi H_t^P \left( \frac{1+\varphi w_{Lt}^{t-1}}{G_t \left( 1+\varphi w_{Lt}^{t-1} \right)} \right) \exp \left( \int_t^\tau \left( g_u^{\pi N} - r_u \right) du \right) d\tau.$$

Further, we know that $r_u = \rho + \theta g_u^C$ 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 $\pi_u^N$ initially grows like $Y_u/N_u$ (and from then on will grow slower), therefore we have that $r_u - g_u^{\pi N} > \rho$. Hence one can write:

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

Since $r_u - g_u^{\pi N} > \rho$, there exists a $\phi > 0$, such that

$$\int_t^{\tilde{t}+T} \exp \left( \int_t^\tau \left( g_u^{\pi N} - r_u \right) du \right) d\tau \leq \int_t^{\tilde{t}+T} e^{-\left(\rho+\phi\right)\left(\tau-t\right)} d\tau$$

$$\leq \frac{1}{\rho + \phi} \left( 1 - e^{-\left(\rho+\phi\right)\left(\tilde{t}+T-t\right)} \right).$$

This allows us to rewrite:

$$\left(\hat{h}_t^A\right)^{1-\kappa} \leq \kappa \eta G_t^{\tilde{\kappa}} \left( \frac{\mu \psi H_t^P \varphi w_{Lt+T}^{t-1}}{\rho} + o \left( \varphi w_{Lt+T}^{t-1} \right) + o \left( e^{-\rho T} \right) \right)$$

Therefore, since $T$ is large and $\varphi w_{Lt+T}^{t-1}$ is small, then $\hat{h}_t^A$ must be small too. In fact, we get that $\hat{h}_t^A = O \left( \left( \varphi w_{Lt+T}^{t-1} \right)^{1-\kappa} \right) + o \left( e^{-\rho T} \right)$.

For any $t \in (0, \hat{t})$, we can then rewrite (47) as

$$\frac{\dot{X}_t}{X_t} = \gamma \psi H^P - \rho - (\theta \psi - \psi + 1) g_t^N + O \left( \left( \varphi w_{Lt+T}^{t-1} \right)^{1-\kappa} \right) + o \left( e^{-\rho T} \right). \quad (123)$$

Using (53) we obtain

$$C_t = Y_t - X_t = (1 + O \left( G_t \varphi w_{Lt}^{t-1} \right)) Y_t.$$
Next \((5)\) and the corresponding equation for high-skill labor demand in production imply:

\[
\frac{L^{NA}}{L^A} = \frac{(1 - G_t) \left(1 + \varphi w_{Lt}^{\sigma - 1}\right)^{\frac{1}{\sigma - 1}}}{G_t} \quad \text{and} \quad \frac{H^{P,NA}}{H^{P,A}} = \frac{(1 - G_t) \left(1 + \varphi w_{Lt}^{\sigma - 1}\right)^{\frac{1}{\sigma}}}{G_t}.
\]

Using \((3)\), we can then write

\[
Y_t = N^{\frac{1}{\sigma - 1}} L^\beta \left( H_t^P \right)^{1 - \beta} \times \left( G_t \left[ 1 + O \left( \varphi w_{Lt}^{\sigma - 1} \right) + \left( O \left( \varphi w_{Lt}^{\sigma - 1} \right) \varphi^2 \frac{Y_t}{L} \right)^{\frac{\theta}{\sigma}} \frac{\frac{\sigma}{\sigma - 1}}{\frac{\frac{\theta}{\sigma}}{\frac{\theta}{\sigma + 1}} \left( 1 + O \left( \varphi w_{Lt}^{\sigma - 1} \right) \right) + 1 - G + O \left( \varphi w_{Lt}^{\sigma - 1} \right) \right]^{\frac{\sigma}{\sigma - 1}} \right).
\]

Note that we have \(w_{Lt} = O \left( Y_t / L \right)\) therefore \(\varphi^2 Y_t / L = O \left( \varphi^2 w_{Lt} \right)\). Therefore

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

From this, using \((8)\), one obtains that high-skill wages obey:

\[
w_{Ht} = \left(1 + O \left( \varphi w_{Lt}^{\sigma - 1} \right) \right) \frac{\sigma - 1}{\sigma} N^{\frac{1}{\sigma - 1}} L^\beta \left( H_t^P \right)^{-\beta},
\]

while for low-skill wages, we get

\[
w_{Lt} = \left(1 + O \left( \varphi w_{Lt}^{\sigma - 1} \right) \right) \frac{\sigma - 1}{\sigma} \beta N^{\frac{1}{\sigma - 1}} L^{\beta - 1} \left( H_t^P \right)^{1 - \beta}.
\]

Therefore using the definition of \(\chi_t\), we obtain that

\[
\chi_t = \left(1 + O \left( \varphi w_{Lt}^{\sigma - 1} \right) \right) \sigma \psi L^\beta \gamma^{(\theta - 1)} \left( H_t^P \right)^{(1 - \beta) \theta + \beta} N^{\frac{1}{\sigma - 1}} L^{\beta - 1} \left( H_t^P \right)^{1 - \beta}.
\]

Differentiating and plugging into \((123)\) and using \((55)\), we get (recalling \((124)\) so that \(d \ln \left(1 + O \left( \varphi w_{Lt}^{\sigma - 1} \right) \right) / dt\) will be of order \(O \left( \varphi w_{Lt}^{\sigma - 1} \right)\) as well).

\[
\frac{\left(1 - \beta\right) \theta + \beta}{H_t^P} \frac{H_t^P}{H_t} = \gamma \psi H_t^P - \rho - \left( \frac{\theta - 1}{\sigma - 1} + 1 \right) \gamma \left( H - H_t^P \right) + O \left( \varphi w_{Lt}^{\sigma - 1} \frac{1}{\frac{\sigma}{\sigma - 1}} \right) + o \left( e^{-\rho T} \right),
\]

we dropped terms in \(\varphi w_{Lt}^{\sigma - 1}\) since there will negligible in front of \(\left( \varphi w_{Lt}^{\sigma - 1} \right)^{\frac{1}{\sigma}}\). The exact counterpart of this system admits a BGP with \(H_t^P\) constant, and as in the Romer
(1990), there is no transitional dynamics. Therefore, here, we must have over the interval 
$(0, \hat{t})$
\[
H_t^P = \left( \frac{\theta - 1}{\sigma} + 1 \right) H + \frac{\theta}{\gamma} + O \left( \left( \varphi w_{L_t+T}^{\epsilon-1} \right)^{\frac{1}{1-\kappa}} \right) + o \left( e^{-\rho T} \right)
\]
and \[g_t^N = \frac{\gamma H \psi - \rho}{\psi + \frac{\theta - 1}{\sigma} + 1} + O \left( \left( \varphi w_{L_t+T}^{\epsilon-1} \right)^{\frac{1}{1-\kappa}} \right) + o \left( e^{-\rho T} \right) \] (126)
which is positive under assumption (25). We then have that for \(N_t\) low, (18) can be solved
as \(G_t = G_0 \exp \left( -\frac{\gamma H \psi - \rho}{\psi + \frac{\theta - 1}{\sigma} + 1} t \right) + O \left( \left( \varphi w_{L_t+T}^{\epsilon-1} \right)^{\frac{1}{1-\kappa}} \right) + o \left( e^{-\rho T} \right) \). This characterizes the
solution during Phase 1 and establishes Proposition 4.A.

### 7.5.3 Transition from the first to the second phase

Here, we prove that if \(\kappa(1 - \beta) + \bar{\kappa} < 1\), \(G_t\) cannot converge toward 0 (that is the first statement of Proposition 4.B).

If \(\bar{\kappa} = 0\), Phase 1 cannot last forever as with positive growth in \(N_t\), \(N_t\) and therefore \(w_{Lt}\) will become large. Since, the Poisson rate is \(\eta \left( \hat{h}_t^A \right)^{\kappa} = O \left( \varphi w_{L_t}^{\epsilon-1} \right). This implies that \(G_t\) must start growing at a positive rate and that we enter the second phase.

When \(\bar{\kappa} > 0\) (and \(G_0 \neq 0\), otherwise automation is impossible), however, whether the
Poisson rate of automation becomes 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 \(G_t w_{L_t}^{\beta(\sigma-1)}\) does not tend towards 0. Then from (122) we obtain
that:
\[
\hat{h}_t^A = \hat{h}_t^A \left( G_t^{\bar{\kappa} - 1} \right)^{\frac{1}{1-\kappa}} \Rightarrow \eta G_t^{\bar{\kappa}} \left( \hat{h}_t^A \right)^{\kappa} = O \left( G_t^{\frac{\bar{\kappa} - 1}{1-\kappa}} \right) \] (127)
Since \(\bar{\kappa} \leq \kappa\), we obtain that the Poisson rate of automation diverges: a contradiction.

Assume now that \(G_t w_{L_t}^{\beta(\sigma-1)}\) does tend towards 0. This ensures that \(w_{Lt} = O \left( N_t^{\frac{1}{\sigma-1}} \right)\).
Moreover, \(\frac{\pi_t^{A} - \pi_t^{N}}{G_t \pi_t^{A} + (1 - G_t) \pi_t^{N}} = O \left( w_{L_t}^{\beta(\sigma-1)} \right)\). Then using this in (122), we obtain
\[
\hat{h}_t^A = O \left( G_t^{\bar{\kappa}} w_{L_t}^{\beta(\sigma-1)} \right)^{\frac{1}{1-\kappa}}.
\]
Note that \(\hat{h}_t^A\) must remain bounded otherwise high-skill labor market clearing is violated.
Therefore, we must have $G_t^\tilde{w}_{Lt}^{\beta(\sigma-1)}$ bounded (which implies that $G_t^\tilde{w}_{Lt}^{\beta(\sigma-1)}$ tends towards 0). Therefore the Poisson rate obeys:

$$\eta_{G_t^\kappa} \left( \tilde{h}_t^A \right)^\kappa = O \left( G_t^{\frac{\kappa}{1-\kappa}} N_t^{\frac{\beta_\kappa}{1-\kappa}} \right)$$

Plugging this in (39) we get:

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

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

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

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

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

$$G_t^\tilde{w}_{Lt}^{\beta(\sigma-1)} = O \left( N_t^{\frac{\beta(1-\kappa)(1-\bar{\kappa})}{1-\kappa-\kappa}} \right),$$

which is indeed declining if $1 - \kappa - \bar{\kappa} < 0$. 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 $\frac{\beta_\kappa}{\kappa+\kappa-1} \leq 1 \iff \kappa(1 - \beta) + \bar{\kappa} \geq 1$.

Alternatively, if $G_t^{\frac{\kappa}{1-\kappa}} N_t^{\frac{\beta_\kappa}{1-\kappa}}$ 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

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

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

Therefore $\kappa(1 - \beta) + \bar{\kappa} < 1$ is a sufficient condition which ensures that the Poisson rate of automation must take off.
**7.5.4 Transitional dynamics in the third phase**

In this appendix, we prove Proposition 4.8 and provide details on the behavior of the economy close to the steady-state. First, we establish:

**Lemma 2.** If $G_t$ is bounded above 0 then $\hat{h}_t^A$ is bounded.

*Proof.* Assume that $G_t$ is bounded above 0 (in fact the analysis of Phase 2 in Appendix 7.5.3 shows that as long as $\kappa (1 - \beta) + \tilde{\kappa} < 1$, 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 a reasoning similar to that in Appendix 7.2.1), 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 $\hat{v}_t$ and $\tilde{c}_t$ are bounded above and below and $\omega_t$ must be bounded above.

Integrating (20), 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 (24), this leads to:

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

Rewriting this expression with the normalized variables and using (52), we get:

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

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

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

for some constant $M$. This ensures that $\frac{V_t^A}{w_{Ht}/N_t}$ must remain bounded, and following (22), $h_t^A$ must be bounded as well.

---

89
Proof of Proposition 4.B Assume now that \( G_t \) has a limit \( G_\infty \) and that \( g_t^N \) also has a positive limit \( g_\infty^N \). Then (39) implies that \( G_\infty < 1 \). Following Proposition (2), we then get that \( w_{Lt} = O \left( N_t \right) \) or \( \omega_t = O \left( 1 \right) \). Therefore, we can rewrite the system as (39).

\[
\dot{h}_t^A = \gamma \frac{1}{\kappa} \left( \hat{h}_t^A \right)^2 - \frac{\eta \kappa G_t^\kappa \left( \hat{h}_t^A \right)^\kappa}{1 - \kappa} \frac{\hat{v}_t^A}{\hat{v}_t} + \eta G_t^\kappa \left( \hat{h}_t^A \right)^\kappa + O \left( n_t \right),
\]

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

Knowing that

\[
H_t^P = \left( \frac{\sigma - 1}{\sigma} \right)^{-\frac{1}{2} \beta \left( \frac{1}{\sigma - 1} \right)} \beta \frac{\mu}{\left( \beta - 1 \right) \sigma \left( 1 - \beta \sigma \right)} \left( G_t \varphi^\mu \right)^{\psi \left( \frac{1}{\sigma - 1} \right)} + O \left( n_t^\mu \right),
\]

(128)

\[
\frac{\hat{v}_t^A}{\hat{v}_t} = \frac{\psi H_t^P \left( \dot{G}_t \right)}{G_t} + O \left( n_t^\mu \right),
\]

(129)

and (55). Using that \( g_t^N \) and \( G_t \) have limits in (39) implies that \( \hat{h}_t^A \) must also have a limit. Using (55), this implies that \( H_t^P \) must also have a limit and therefore using (128) that \( \chi_t \) must have a limit. In other words, the equilibrium path tends toward the steady-state \( \left( \hat{h}_t^A, G_\infty, \chi_\infty \right) \) with \( n_t \to 0 \) defined in Proposition 3.

Behavior close to the steady-state. In this steady-state, using (59) we get

\[
g_t^N = \frac{1}{\left( \theta - 1 \right) \psi + 1} \left( \gamma \frac{1 - \kappa}{\kappa} \hat{h}_t^A - \rho \right).
\]

Therefore in the third phase, we obtain that \( N_t \) grows at rate \( g_t^N = g_{N_t}^* + o \left( 1 \right) \), that the share of automated product obeys \( G_t = G_\infty + o \left( 1 \right) \), with the mass of high-skill workers in automation given by \( H_t^A = \left( 1 - G_\infty \right) \hat{h}_t^A + o \left( 1 \right) \) and the mass of high-skill workers in production given by \( H_t^P = H_\infty^P + o \left( 1 \right) \), with \( H_\infty^P \) given by (60). Using (50), we obtain
that wages obey:

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

and

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

Using (129), the profit made by an automated firm are given by

\[ \pi^A_t = \frac{1}{\sigma} \left( \frac{\sigma - 1}{\sigma} \beta^\beta \right) \frac{1}{1-\beta} H^* (G^*)^{\psi-1} N_t^{\psi-1} + o \left( N_t^{\psi-1} \right), \]

while the profits made by a non-automated firm \( \pi^N_t \) are negligible in front of \( \pi^A_t \). Using (40), the value of an automated firm is then simply given by:

\[ V^A_t = \frac{\pi^A_t}{r^* - (\psi - 1) g^{N*}} + o \left( N_t^{\psi-1} \right), \]

where

\[ r^* = \rho + \theta \psi g^{N*} \]

is the steady-state interest rate. Following (41) and (42), the normalized value of a non-automated firm obeys:

\[ (r_t - (\psi - 1) g_t^N) \tilde{V}^N_t = \tilde{\pi}^N_t + (1 - \kappa) \eta G^t \tilde{h}^A_t \left( \tilde{V}^A_t - \tilde{V}^N_t \right) + \tilde{V}^N_t. \]

Therefore, one gets that for large \( N_t \),

\[ V^N_t = \frac{(1 - \kappa) \eta G^N \tilde{h}^{A*}}{r^* - (\psi - 1) g^{N*} + (1 - \kappa) \eta G^N \tilde{h}^{A*}} V^A_t + o \left( N_t^{\psi-1} \right). \]

7.5.5 Comparing the growth rate in the number of products in Phase 1 and Phase 3.

Use (130), (132) and (131) to get that in steady-state:

\[ \tilde{V}^{N*} = f \tilde{V}^{A*} \]
with

\[
\begin{align*}
    f &= \frac{(1 - \kappa) \eta G^* \hat{h}^A*}{\rho + (\psi (\theta - 1) + 1) g_N^* + (1 - \kappa) \eta G^* \hat{h}^A*}, \\
    \hat{V}^A* &= \frac{\hat{\pi}_t^A}{\rho + (\psi (\theta - 1) + 1) g_N^*}.
\end{align*}
\]

(133)

Using \(\hat{V}^N* = \hat{\sigma}^*/\gamma\) and (129), we then obtain:

\[
\frac{1}{f} \frac{\psi H^P*}{\rho + (\psi (\theta - 1) + 1) g_N^*} = \frac{1}{g}. \gamma
\]

Rearranging terms, this leads to

\[
g^N* = \frac{\frac{\psi H^P*}{\rho + (\psi (\theta - 1) + 1) g_N^*} \psi H^P* - \rho}{\frac{1}{\sigma} \psi + \frac{\psi}{1 - 1 (1 - \beta)} + 1},
\]

(134)

while from (126) the growth rate in the first period is approximately given by

\[
g^N1 = \frac{\gamma H^P* - \rho}{\psi + \frac{\psi}{1 - 1} + 1}. \]

(135)

The two expressions differ by three terms: In the numerator, \(H - (1 - G^*) \hat{h}^A*\) in (134) replaces \(H\) in (135), since some high-skill workers are hired to automate in Phase 3, 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, \(\frac{\theta - 1}{\sigma - 1}\) in (134) replaces \(\frac{\theta - 1}{\sigma - 1}\) in (135), this reflects the fact that the growth rate in the number of products has a larger impact on the economy growth rate in phase 3 than in phase 1, which in return reduces the present value of an automated firm (it increases the effective interest rate). This also pushes toward \(g^N* < g^N1\). Finally the term \(f/G^*\) in (134) does not exist in (135). Note that \(\partial g^N*/(\partial f/G^*) > 0\), so that this term reflects two different forces. On one hand in phase 3, 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 (39), (58) and (133), we get

\[
\frac{f}{G^*} < 1 \iff \frac{1}{G} (1 - \kappa) \eta G^* \hat{h}^A < \rho + (\psi (\theta - 1) + 1) g^*_N + (1 - \kappa) \eta G^* \hat{h}^A \\
\iff \frac{1 - G}{G} (1 - \kappa) \eta G^* \hat{h}^A < \rho + (\psi (\theta - 1) + 1) g^*_N, \\
\iff (1 - \kappa) g^*_N < \rho + (\psi (\theta - 1) + 1) g^*_N.
\]

Since \(\theta \geq 1\) and \(\kappa < 1\), this equation necessarily holds. Therefore \(f/G^* < 1\) so that, interpreting the value of a new firm as the discounted flow of a “net profit flow” \(f\hat{\pi}_t^A\), we get that the profit flow of new firms in the asymptotic steady-state is a lower fraction of total output than the profit flow of new firms in the first phase, ensuring that \(g^N > g^N_1\). This establishes Proposition 5.

7.5.6 Comparative statics

In this section, we prove Proposition 6. 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 \(f (g^N) = 1\) through (62), where \(G^*, \hat{h}^A,\) 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 \(g^N\).

**Comparative static with respect to \(\gamma\):** (59) implies that \(\hat{h}^A\) is inversely proportional to \(\gamma\) (for given \(g^N\)). Formally, we have:

\[
\frac{\partial \hat{h}^A}{\partial \gamma} = -\frac{\hat{h}^A}{\gamma}. \tag{136}
\]

Differentiating (58) and using (136) leads to:

\[
\frac{\partial G^*}{\partial \gamma} = \frac{-\kappa g^N G^*}{\gamma \left(\eta (G^*) \hat{\kappa} \left(\hat{h}^A\right)^\kappa + (1 - \hat{\kappa}) g^N\right)} \tag{137}
\]
so that for a given $g^{N*}$, $G^*$ is also decreasing in $\gamma$. Using (60), (136) and (137), we get:

$$\frac{\partial H^P^*}{\partial \gamma} = \frac{1}{\gamma} \left( \frac{g^{N*}}{\gamma} + \frac{(1 - G^*) (1 - \kappa) \hat{h}^{A*}}{(1 - \kappa) \kappa \hat{h}^{A*} + (1 - \kappa) \gamma} \right) > 0$$

so that $H^P^*$ is increasing in $\gamma$. Note that $f$, defined in (62), can be rewritten as

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

which shows that $f$ is decreasing in $\gamma$ for a given $g^{N*}$ ($H^P^*$ is increasing, $G^*$ and $\hat{h}^{A*}$ are decreasing, and $\gamma \hat{h}^{A*}$ is constant). Since $f$ is increasing in $g^{N*}$ at the equilibrium value, (62) implies that $g^{N*}$ increases in $\gamma$.

**Comparative static with respect to $\eta$.** For given $g^{N*}$, (59) implies that $\hat{h}^{A*}$ does not depend on $\eta$. Differentiating (58), we get:

$$\frac{\partial \ln G^*}{\partial \ln \eta} = \frac{g^{N*}}{\eta (G^*)^{\kappa} \left( \hat{h}^{A*} \right)^{\kappa} + (1 - \kappa) g^{N*}}, \quad (138)$$

so for given $g^{N*}$, $G^*$ increases in $\eta$. (60) implies then that

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

Using this equation together with (138) and (62), we obtain:

$$\frac{\partial \ln f}{\partial \ln \eta} = \left\{ \frac{g^{N*}}{\eta (G^*)^{\kappa} \left( \hat{h}^{A*} \right)^{\kappa} + (1 - \kappa) g^{N*}} \left( 1 - \frac{G^* \hat{h}^{A*}}{H^P^*} - \kappa \frac{1}{\kappa \eta (G^*)^{\kappa} \left( \hat{h}^{A*} \right)^{1 - \kappa} + \frac{1}{\gamma}} \left( \hat{h}^{A*} \right)^{1 - \kappa} + \frac{1}{\gamma} \right) \right\}.$$
Using (59), we can rewrite this as:

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

so that \( f \) is decreasing in \( \eta \). This implies that \( g^{N*} \) must be increasing in \( \eta \). Since \( \hat{h}^A \) only depends on \( \eta \) through \( g^{N*} \), we also get that \( \hat{h}^A \) increases in \( \eta \).

### 7.6 Simulation technique

In the following we describe the simulation techniques employed for the baseline model presented in [2]. The approach for the extensions follow straightforwardly. Let \( x_t \equiv (n_t, G_t, \hat{h}_t^A, \chi_t, \omega_t) \) and note that equation (56) defines \( \omega \) implicitly. We can therefore write equations (38), (39), (46) and (47) as a system of autonomous differential equations

\[
\dot{n}_t, \quad \dot{G}_t, \quad \dot{\hat{h}}_t^A, \quad \dot{\chi}_t = F(x_t)
\]

with initial conditions on state variables as \((n_0, G_0)\) and an auxiliary equation of \( \dot{\omega}_t = 0 \).

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:

\[
s_t = \left( \frac{n_{t+1} - n_t}{dt}, \frac{G_{t+1} - G_t}{dt}, \frac{\hat{h}_{t+1}^A - \hat{h}_t^A}{dt}, \frac{\chi_{t+1} - \chi_t}{dt} \right) - F((x_t + x_{t+1})/2), \quad t = \{1,...T - 1\}
\]

with \( T \) corresponding errors for \( \omega_t \):

\[
s^\omega_t = \omega_t - \vartheta(x_t), \quad t = \{1,...,T\}.
\]

Following Proposition 3 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 \( \hat{h}_T^A \) and \( \chi_T \) accordingly. Together with the initial conditions \((n_1 = n_{\text{start}} \text{ and } G_1 = G_{\text{start}})\) this leads to a vector of errors:

\[
s_T \equiv (n_1 - n_{\text{start}}, G_1 - G_{\text{start}}, \hat{h}_T^A - \hat{h}^A, \chi_T - \chi^*)'.
\]
Letting \( x = \{x_t\}_{t=1}^T \), we then stack errors to get a vector, \( S(x) \), of length \( 5T \) and solve the following problem:

\[
\hat{x} = \text{argmin}_x S(x)'WS(x),
\]

for a \( 5T \times 5T \) diagonal weighting matrix, \( W \), and the \( 5T \) vector \( x \). For \( dt \to 0 \) and \( T \to \infty \) \( S(x)'WS(x) \to 0 \). For the simulations we set \( dt = 2 \) and \( T = 2000 \). We accept the solution when \( S(\hat{x})'WS(\hat{x}) < 10^{-7} \), 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 \( \{\hat{x}_t, \hat{\omega}_t\}_{t=1}^T \) in hand, it is straightforward to find remaining predicted values.

7.7 Complements on simulation

7.7.1 Additional results on the baseline simulation

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.

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 7.A), which is in line with a stable ratio of total R&D expenses over GDP across the three phases (Figure 7.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 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. As claimed in the text, 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, which still need to be automated (foll value.

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? 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.

More specifically we can write \( w_{Lt} = f(N_t, G_t, H_t^P) \), using equations (10) and (11). Differentiating with respect to time and using equation (39) gives:

\[
\dot{w}_{Lt} = \left( \frac{N_t}{w_{Lt}} \frac{\partial f}{\partial N} - \frac{G_t}{w_{Lt}} \frac{\partial f}{\partial G} \right) \gamma H_t^D + \frac{1}{w_{Lt}} \frac{\partial f}{\partial G} \eta G_t^\alpha (1 - G_t) (\hat{h}_t^A)^\kappa + \frac{1}{w_{Lt}} \frac{\partial f}{\partial H} \dot{H}_t^P.
\]

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 \( w_{Ht} \).
This should not be interpreted as “automation being harmful” to low-skill workers in general. In fact, as we demonstrate in Section 3.5, an increase in the effectiveness of the automation technology, $\eta$, 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).

### 7.7.2 The effect of the innovation parameters

Figure 14 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 $\eta$ 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 $\eta$ 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, $\gamma$, 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.
7.7.3 Technological shocks

Figure 15 shows the effect of a technological shock in the form of an exogenous change in $G_t$ in year 150 (for the same parameters as in the baseline case). Panel A shows that after an increase in $G$, the mass of high-skill workers in automation innovation declines, this guarantees that the share of automated products goes back to its initial path (panel B). The skill premium is reduced by the shock but thereafter increases relative to the baseline level (panel C). It does not converge back to exactly the same level as before, however, because the shock to $G_t$ and the following decrease in $H_t^A$ implies that for some time more high-skill workers undertake horizontal innovation which increases a bit the number of automated products relative to the baseline.

7.7.4 Systematic comparative statics

In this section we carry a systematic comparative exercise with respect to the parameters of the model, namely $\sigma, \epsilon, \beta, \rho, \theta, \phi, \eta, \kappa, \tilde{\kappa}, \gamma, H/L$ (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 16: 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 16.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 16.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 µ < 1), a higher elasticity reduces the growth of low-skill wages and increases that of high-skill wages during Phase 2. Figures 16.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 17: Comparative statics with respect to the discount rate ($\rho$), the inverse elasticity of intertemporal substitution ($\theta$) and the productivity of machines ($\tilde{\phi}$).

Figures 17.A,B,C show that a higher discount rate $\rho$ 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 $\theta$) has a similar effect on the economy’s growth rate (Figures 17.D,E,F). Figures 17.G,H,I show that the productivity of machines ($\tilde{\phi}$) only affects the timing of Phase 2 (Phase 2 occurs sooner when machines are more productive).
Figure 18: Comparative statics with respect to the automation productivity ($\eta$), the concavity of the automation technology ($\kappa$) and the automation externality ($\tilde{\kappa}$)

The comparative statics with respect to the automation technology shown in Figures 18.A,B,C follow the pattern described in the text. A less concave automation technology (higher $\kappa$) 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 18.D,E,F. The role of the automation externality has already been discussed in the text, Figures 18.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.
Figures 19.A,B,C show the impact of the horizontal innovation parameter $\gamma$, which was already discussed in the text. Figures 19.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 20.A,B,C show that a higher initial number of products simply advance the entire evolution of the economy. Figures 20.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.

7.8 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: $N^A_t$ (respectively $N^N_t$) denotes the mass of automated (respectively non-automated) firms, $L^A_t$ (respectively $L^N_t$) is the mass of low-skill workers hired in automated (respectively non-automated) firms, and $H^{P,A}_t$ (respectively $H^{P,N}_t$) 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
Figure 20: Comparative statics with respect to the initial number of products $N_0$ and the initial share of automated products $G_0$

Lagrange multipliers next to each constraint):

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

such that

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

with

$$F \equiv \left( (N_t^A)^{\frac{1}{\sigma}} \left( \left( \frac{2}{\sigma} X_t^{\frac{\sigma-1}{\sigma}} + \left( L_t^A \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \left( H_t^{P,A} \right)^{1-\beta} \right)^{\frac{\sigma-1}{\sigma}} \right),$$

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

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

$$\bar{\zeta}_t: N_t^N = \gamma (N_t^A + N_t^N) H_t^D - \eta (N_t^A)^{\frac{\kappa}{1-\kappa}} (N_t^N + N_t^A)^{\kappa-\frac{\kappa}{1-\kappa}} (H_t^A)^{\frac{\kappa}{1-\kappa}} (N_t^N)^{1-\kappa} ,$$

$$\bar{\xi}_t: N_t^A = \eta (N_t^A)^{\frac{\kappa}{1-\kappa}} (N_t^N + N_t^A)^{\kappa-\frac{\kappa}{1-\kappa}} (H_t^A)^{\frac{\kappa}{1-\kappa}} (N_t^N)^{1-\kappa} ,$$
The first order condition with respect to $C_t$ gives

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

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

The first order conditions with respect to $X_t$ implies that

$$\frac{\partial F}{\partial X_t} = 1, \quad (139)$$

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

$$w_{Lt} = \frac{\partial F}{\partial L^A_t} = \frac{\partial F}{\partial L^N_t} \text{ and } w_{Ht} = \frac{\partial F}{\partial H^{P,A}_t} = \frac{\partial F}{\partial H^{P,N}_t}, \quad (140)$$

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 $H^P_t$, 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 $N^N_t$ and $N^A_t$ are given by:

$$\rho \tilde{\xi}_t - \tilde{\zeta}_t = \tilde{\lambda}_t \frac{\partial F}{\partial N^N_t} + \tilde{\zeta}_t \gamma H^D_t + \left( \tilde{\xi}_t - \tilde{\zeta}_t \right) \eta \left( H^A_t \right)^{\kappa} \left( N^N_t \right)^{-\kappa} \left( \left( 1 - \kappa \right) N^N_t + \left( 1 - \kappa \right) N^A_t \right) \left( N^N_t + N^A_t \right)^{\kappa-\tilde{\kappa}-1}, \quad (141)$$

$$\rho \tilde{\xi}_t - \tilde{\zeta}_t = \tilde{\lambda}_t \frac{\partial F}{\partial N^A_t} + \tilde{\zeta}_t \gamma H^D_t + \left( \tilde{\xi}_t - \tilde{\zeta}_t \right) \eta \left( H^A_t \right)^{\kappa} \left( \kappa N^N_t + N^A_t \right) \left( N^N_t + N^A_t \right)^{\kappa-\tilde{\kappa}-1}. \quad (142)$$

Interestingly, $\frac{\partial F}{\partial N^N_t}$ and $\frac{\partial F}{\partial N^A_t}$ 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

\[ \pi_t^N = \frac{\partial F_t}{\partial N_t^N} \quad \text{and} \quad \pi_t^A = \frac{\partial F_t}{\partial N_t^A} \]

Further the (shadow) interest rate is given by \( r_t = \rho + \theta \frac{G_t}{C_t} = \rho - \frac{\lambda_t}{G_t} \). Using that \( H_t^A = (1 - G_t) N_t h_t^A \), we can rewrite (141) and (142) as:

\[ r_t \xi_t = \pi_t^N + \xi_t g_t^N + (\xi_t - \zeta_t) \eta (G_t) \tilde{\kappa} N_t^\kappa (h_t^A)^\kappa ((1 - \tilde{\kappa}) (1 - G_t) + (1 - \kappa) G_t) + \dot{\zeta}_t, \quad (143) \]

\[ r_t \zeta_t = \pi_t^A + \zeta_t g_t^N + (\xi_t - \zeta_t) \eta (G_t) \tilde{\kappa} N_t^\kappa (h_t^A)^\kappa (1 - G_t) \left( \tilde{\kappa} \frac{1 - G_t}{G_t} + \kappa \right) + \dot{\zeta}_t. \quad (144) \]

These expressions parallel equations (20) and (21) in the paper. The rental social value of a non-automated firm \( (r_t \xi_t) \) 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 \( \gamma N_t \)), 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 \( H_t^D \geq 0 \)):

\[ w_{Ht} \geq \zeta_t \gamma N_t, \quad (145) \]

with equality when \( H_t^D > 0 \). This equation is the counterpart of (23) 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:

\[ w_{Ht} = (\xi_t - \zeta_t) \kappa \eta (G_t) \tilde{\kappa} N_t^\kappa (h_t^A)^{\kappa - 1} . \quad (146) \]
This equation is the counterpart of (22) 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 \( \left( n_t, G_t, \hat{h}_t^A, \chi_t \right) \) which characterizes the solution (when there is positive growth). Equations (38) and (39) still hold, while equations (46) and (47) are replaced with

\[
\frac{\hat{A}^A}{\hat{h}_t} = \frac{\gamma \hat{h}_t^A}{1-\kappa} \left( \omega_t n_t \left( \varphi + \left( \omega_t n_t \right)^{-\beta} \right)^\mu \right) \left( \frac{1}{1-G_t} \right) \hat{h}_t^A + \frac{\mu \hat{A}^A}{\hat{v}_t} + \frac{1-\kappa+(\kappa-\check{\kappa})(1-G_t)}{\kappa} \left( \hat{h}_t^A \right)^{\kappa+1} + \frac{\gamma t}{1-\kappa} g_t^N \hat{h}_t^A,
\]

\[
\dot{\chi}_t = \chi_t \left( \gamma t n_t \left( \varphi + \left( \omega_t n_t \right)^{-\beta} \right)^\mu \right) \left( \frac{1}{1-G_t} \right) \hat{h}_t^A - \rho - (\theta - 1) \psi g_t^N.
\]

\( g_t^N \) is still given by (55), \( \frac{\hat{A}^A}{\hat{v}_t} \), \( H_t^P \) and \( \omega_t \) are now given by

\[
\frac{\hat{A}^A}{\hat{v}_t} = \frac{\psi \left( \varphi + \left( \omega_t n_t \right)^{-\beta} \right)^\mu}{G_t \left( \varphi + \left( \omega_t n_t \right)^{-\beta} \right)^\mu + (1-G_t) \omega_t n_t},
\]

\[
H_t^P = \frac{\left( 1 - \beta \right) ^{\frac{1}{\beta - 1}} \beta t \bar{\pi}^{\bar{\pi}} \left( \hat{A}^A \right) \chi_t \left( \varphi + \left( \omega_t n_t \right)^{\frac{1}{\beta}} \right)^\mu + (1-G_t) \omega_t n_t \right) \psi \left( \frac{1}{\beta} - 1 \right) + 1}{G_t \left( (1-\beta) \varphi + (\omega_t n_t)^{\frac{1}{\beta}} \right) \left( \varphi + (\omega_t n_t)^{\frac{1}{\beta}} \right)^{\mu-1} + (1-G_t) \omega_t n_t},
\]

\[
\omega_t = \left( \beta \frac{H_t^P}{\hat{L}} \left( G_t \left( \varphi + (\omega_t n_t)^{\frac{1}{\beta}} \right)^{\mu-1} (\omega_t n_t)^{\frac{1}{\beta}} + (1-G_t) \right) \right) \frac{\beta \left( \frac{1}{\beta} - 1 \right) + 1}{\beta \left( \frac{1}{\beta} - 1 \right) + 1 + \beta \left( \frac{1}{\beta} - 1 \right)},
\]

which replace (52), (54) and (56).

One can then solve for a steady state of this system with \( G^* > 0 \) (and \( g^N > 0 \) so
that \( n^* = 0 \). (58) and (60) still apply, but (59) is replaced with

\[
\tilde{h}^{A^*} = \frac{\kappa + (\theta - 1) \psi g^{N*}}{\gamma 1 - \kappa + (1 - G^*) (\kappa - \tilde{\kappa})},
\]

(148) and (62) with

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

which is obtained by fixing \( \tilde{h}_t = 0 \) in (147) using (58) and (148). For \( g^{N*} \) large enough (but finite—and, in particular smaller than \( \gamma H \)), \( H^{P*} \) is arbitrarily small, while for the same value \( G^* \) and \( \tilde{h}^{A^*} \) are bounded below and above. As before, this establishes that for \( g^{N*} \) large enough \( f^{sp} (g^{N*}) > 1 \). Furthermore \( f^{sp} (0) = f (0) \), therefore condition 25 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 \( (s^A_t) \) and in horizontal innovation \( (s^H_t) \). Let us consider such an equilibrium and introduce the notations \( \Omega^A_t \equiv 1 - s^A_t \) and \( \Omega^H_t \) similarly defined. In this situation, the law of motion for the private value of an automated firm, \( V_t^A \), is still given by (20), for a non-automated firm it obeys:

\[
\eta (G_t)^{\kappa} \left( h_t^A \right)^{\kappa} \left( V_t^A - V_t^N \right) + \dot{V}_t^N = \pi_t^N - \Omega^A_t w_H t h_t + \eta (G_t)^{\kappa} N_t^\kappa \left( h_t^A \right)^{\kappa} \left( V_t^A - V_t^N \right) + \dot{V}_t^N, \tag{149}
\]

instead of (21), the first-order condition for automation is given by:

\[
\kappa \eta (G_t)^{\kappa} \left( h_t^A \right)^{\kappa-1} \left( V_t^A - V_t^L \right) = \Omega_t^A w_H t, \tag{150}
\]
instead of (22), while the free entry condition, when \( g_t^N > 0 \), is given by

\[
\gamma N_t V_t^N = \Omega_t^H w_{Ht},
\]

(151)

instead of (23). For \( \Omega_t^A \) and \( \Omega_t^H \) to decentralize the optimal allocation it must be that these 4 equations hold together with (143), (144), (145) and (146).

Using (145) and (151), we then get that \( \Omega_t^H \) must satisfy

\[
\Omega_t^H \zeta_t = V_t^N,
\]

(152)

similarly, using (146) and (150), we get

\[
\Omega_t^A (\xi_t - \zeta_t) = V_t^A - V_t^L.
\]

(153)

Plugging (152) and (153) in (149), we get that

\[
r_t \zeta_t = \pi_t^N \Omega_t^H - \frac{\Omega_t^A}{\Omega_t^H} w_{Ht} \hat{h}_t + \eta (G_t)^\kappa N_t^\kappa (h_t^A)^\kappa \Omega_t^A (\xi_t - \zeta_t) + \frac{\Omega_t^H}{\Omega_t^H} \zeta_t + \dot{\zeta}_t.
\]

(154)

Similarly, using (153) and the difference between (20) and (149) gives:

\[
r_t (\xi_t - \zeta_t) = \frac{\pi_t^A - \pi_t^N}{\Omega_t^A} + w_{Ht} \hat{h}_t - \eta (G_t)^\kappa N_t^\kappa (h_t^A)^\kappa (\xi_t - \zeta_t) + \frac{\Omega_t^A}{\Omega_t^A} (\xi_t - \zeta_t) + \dot{\xi}_t - \dot{\zeta}_t.
\]

(155)

Combining (154) with (143), using (146) and (145) and the definition of \( \Omega_t^A \) and \( \Omega_t^H \), we get:

\[
\frac{\gamma \hat{h}_t^A}{\nu_t} \left( (1 - s_t^A) (1 - \kappa) + (1 - s_t^H) (\kappa (1 - G_t) + \kappa G_t - 1) \right).
\]

(156)

Similarly combining (155) with the difference between (144) and (143) and using (145) gives:

\[
\frac{\gamma \hat{h}_t^A}{\nu_t} \left( (1 - s_t^A) (1 - \kappa) + (1 - s_t^H) (\kappa (1 - G_t) + \kappa G_t - 1) \right).
\]

(157)
Therefore, in steady state, we have

\[ s_\infty^A = \frac{\tilde{\kappa}h_\infty^A (1 - G_\infty)}{\kappa \psi H_\infty^P + \tilde{\kappa}h_\infty^A (1 - G_\infty)} \geq 0. \]

Note from (157) 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 \( \tilde{\kappa} \neq 0 \)). Furthermore, if \( \tilde{\kappa} = 0 \), \( s_t^A = 0 \) 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).

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

\[ s_\infty^H = 1 - \frac{\gamma \tilde{h}_\infty^A (1 - \kappa) (1 - s_\infty^A)}{\kappa g_\infty^N + \gamma \tilde{h}_\infty^A (1 - \tilde{\kappa} (1 - G_\infty) - \kappa G_\infty)}. \]

In addition, knowing that \( s_t^A \geq 0 \), imposes that \( s_t^H > 0 \)—as \( s_t^H < 0 \) would lead to \( s_t < 0 \).

### 7.8.4 Translational dynamics for the social planner case

Figure 21 plots the transitional dynamics for the optimal allocation in our baseline case (with \( \tilde{\kappa} = 0 \)) and in the case where \( \tilde{\kappa} > 0 \) analyzed in Figure 8. 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 7.A and Figure 8.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 \( \tilde{\kappa} = 0 \) (without the automation externality), the subsidy to automation is 0, while when \( \tilde{\kappa} > 0 \) 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.
7.9 Alternative model with automation at the entry-stage

To highlight that the evolution of the economy through three phases does not depend on our assumption that new products are born non-automated, we present in this section a model where, instead, we assume that automation can only take place at the entry stage. That is, when a new firm is born, it can hire \( h_t^A \) workers to automate it, in which case it is successful with probability \( \min (\eta (N_t h_t^A)^\kappa, 1) \) (we abstract from the automation externality for simplicity). Ex-ante a firm does not know whether it will succeed or not, therefore, the free-entry condition can now be written as

\[
\begin{align*}
  w_{HT} & \geq \gamma N_t V_t,
\end{align*}
\]

where

\[
V_t = \min \left( \eta (N_t h_t^A)^\kappa, 1 \right) V_t^A + \left( 1 - \min \left( \eta (N_t h_t^A)^\kappa, 1 \right) \right) V_t^N - w_{HT} h_t^A.
\]

is the expected value of a new firm. Since we used similar functional forms we have that \( h_t^A \) obeys (22) unless \( \kappa \eta^\frac{1}{2} N_t (V_t^A - V_t^L) > w_{HT} \), in which case \( N_t h_t^A = \eta^{-\frac{1}{2}} \). Afterward a
firm never becomes automated so that the law of motion for the value of an automated
and a non-automated firms both follow (20). In addition, the law of motion for \( G_t \) is now given by
\[
\dot{G}_t = g_t^N \left( \eta \left( N_t h_t^A \right)^\kappa - G_t \right).
\]
The resolution of the model follows the same steps as in the baseline case, and under
the appropriate condition on the discount rate, there exists an asymptotic steady state
with \( g_t^N > 0 \). It is still the case that a higher level of \( G_t \) requires innovation to be more
directed toward automation (i.e. a larger level for \( \eta \left( N_t h_t^A \right)^\kappa \)) in order to ensure that \( G_t \)
increases. As a result, even though only new products can be automated, it is still the
case that the skill premium need not increase the fastest when technology is the most
directed toward automation.

An important difference is that \( G^* \) may be equal to 1 since all new products may
choose to be automated in steady state. In fact, one can derive that \( \hat{h}^A^* = \min \left( \eta^{-\frac{1}{\kappa}}, \frac{\kappa}{1-\frac{1}{\gamma}} \right) \).
Therefore \( G^* < 1 \), if and only if \( \eta \left( \frac{\kappa}{1-\frac{1}{\gamma}} \right)^\kappa < 1 \). When \( G^* < 1 \), we will have that
\( G_\infty = G^* < 1 \), so that, following Proposition 2,
\[
g^w_L = \frac{1}{1 + \beta \left( \sigma - 1 \right)} g^w_H.
\]
On the contrary, if \( G^* = 1 \), then \( G_\infty = 1 \), and following Proposition 2 we get that
\[
g^w_L = g^w_H / \epsilon.
\]

Figure 22 draws the transitional dynamics for the same parameters as in the baseline
case (even though the automation technology parameters have a different meaning here).
These parameters satisfy \( \eta \left( \frac{\kappa}{1-\frac{1}{\gamma}} \right)^\kappa < 1 \), and the figure shows that the economy goes
through three phases as in our baseline model.

### 7.10 An endogenous supply response in the skill distribution

We present here an extension of the baseline model with an endogenous supply re-
response in the skill distribution. Specifically, let there be a unit mass of heterogeneous
individuals, indexed by \( j \in [0,1] \) each endowed with \( lH \) units of low-skill labor and
\( \Gamma (j) = H^{1+q} j^{1/q} \) units of high-skill labor (the important assumption here is the ex-
istence 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 (11) which still holds. (10) also holds with \( L_t \) replacing \( L \) and knowing that \( H_t^P \) obeys (19) 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 \( j \in (0, \bar{j}_t) \) supply low-skill labor and workers \( j \in (\bar{j}_t, 1) \) supply high-skill labor. As a result, we get that the total mass of low-skill labor is:

\[
L_t = l \bar{H} \bar{j}_t ,
\]

and the mass of high-skill labor is

\[
H_t = \bar{H} \left( 1 - \frac{1 + \frac{q}{\bar{H} w_{Lt}}}{} \right) \leq \bar{H}.
\]

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

\[
\bar{j}_t = \left( \frac{q}{1 + q w_{Ht}} \right) \frac{l w_{Lt}}{\bar{H} w_{Ht}}.
\]

\( \bar{j}_t \) decreases as the skill premium increases and \( q \) measures the elasticity of \( \bar{j}_t \) 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 10.** 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 \(G_t, g_t^N, H_t^P\) all admit limits \(G_\infty, g_\infty^N\) and \(H_\infty^P\) 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:

\[
g^{wH}_\infty = g^Y_\infty = g^N_\infty \quad \text{and} \quad g^{wL}_\infty = \frac{1 + q}{1 + q + \beta(\sigma - 1)} g^Y_\infty.
\]

**(161)**

**Proof.** We consider processes \((N, G_t, H_t^P)\) such that \(g_t^N, G_t\) and \(H_t^P\) admit strictly positive limits. Plugging (160) and (158) in (10), we get:

\[
\frac{w_{Ht}}{w_{Lt}} = l \left( \frac{1 - \beta}{\beta} \frac{H}{H_t^P} \right)^q \frac{G_t + (1 - G_t) (1 + \varphi w_{Lt}^{-1})^{-\mu}}{G_t (1 + \varphi w_{Lt}^{-1})^{-1} + (1 - G_t) (1 + \varphi w_{Lt}^{-1})^{-\mu}} \right)^{\frac{1}{1+q}},
\]

which together with (11) determines \(w_{Ht}\) and \(w_{Lt}\) for given \((N_t, G_t, H_t^P)\). From then on the reasoning follows that of Appendix 7.2.1. First, we derive that \(w_{L\infty} > 0\), such that \(g^{wL}_\infty = \psi g^N_\infty\), and that we must have \(g^{wL}_\infty < g^{wH}_\infty\), such that \(\tilde{J}_\infty = 0\). Second, we study the asymptotic behavior of \(w_{Lt}\) both when \(\epsilon < \infty\) and when \(\epsilon = \infty\).

**Case with \(\epsilon < \infty\).** Plugging (162) in (11) gives \(w_{Lt}\) in function of \(N_t, G_t\) and \(H_t^P\):

\[
w_{Lt} = \frac{\sigma-1}{\sigma} \beta \left( \frac{H}{H_t^P} \right)^{1+\frac{1}{1+q}} \left( 1 - \beta \right)^{1+\frac{1}{1+q}} \left( \frac{G_t + (1-G_t) (1+\varphi w_{Lt}^{-1})^{-\mu}}{G_t (1+\varphi w_{Lt}^{-1})^{-1} + (1-G_t) (1+\varphi w_{Lt}^{-1})^{-\mu}} \right)^{\frac{1}{1+q}} N^{\frac{1}{\sigma-1}},
\]

which replaces (72). It is direct that when \(G_\infty < 1\), we obtain (161). In this case, we further have

\[
g^\tilde{J}_\infty = q (g^{wL}_\infty - g^{wH}_\infty) = -\frac{q \beta (\sigma - 1)}{1 + q + \beta (\sigma - 1)} g^{GDP}_\infty.
\]

**(164)**

**Case with \(\epsilon = \infty\).** In this case, (163) becomes

\[
w_{Lt} = \frac{\sigma-1}{\sigma} \beta \left( \frac{H}{H_t^P} \right)^{1-\beta} \left( \frac{1+q}{q} \right)^{1+\frac{1}{1+q}} \left( \frac{G_t + (1-G_t) (1+\varphi w_{Lt}^{-1})^{-\mu}}{G_t (1+\varphi w_{Lt}^{-1})^{-1} + (1-G_t) (1+\varphi w_{Lt}^{-1})^{-\mu}} \right)^{\frac{1}{1+q}} N^{\frac{1}{\sigma-1}} (1-G_t)^{\frac{1}{1+q}},
\]

if \(w_{Lt} > \varphi^{-1}\),

115
Once again, following the steps of Appendix 7.2.1 we get that if $G_\infty < 1$, (161) applies (and accordingly we also get (164)).

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.

We can then derive the dynamic system as in the baseline model (see details below in Appendix 7.10.1). Proposition 3 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 2, $\bar{H} = 1/3$ (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 7, 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.

7.10.1 Details on the dynamic system

It is convenient to redefine $n_t \equiv N_t^{(1-\beta)\frac{1+q}{1+q+\beta(\sigma-1)}}$, we can then write the entire dynamic system as a system of differential equations in $\left(n_t, G_t, \hat{h}^A_t, \chi_t\right)$ with two auxiliary variables $\omega_t$ and $\hat{j}_t \equiv \bar{j} t n_t^{-\frac{q}{1+q}}$. Equations (38) is now given by

$$\dot{n}_t = -\frac{\beta}{1-\beta} \frac{1+q}{1+q+\beta(\sigma-1)} g^N_t n_t.$$
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.

\( (39), (46), (47), (52), (54) \) still apply and equation \( (55) \) as well provided that \( H \) is replaced by \( H_t \) given by (159). \( \omega_t \) is implicitly defined by:

\[
\omega_t = \left( \frac{1+q}{1-\beta} \right)^{\frac{1}{1+q}} \beta^{\frac{1+q}{1-\beta}} \left( \frac{1}{1-\beta} \right)^{\frac{1+q}{q}} \frac{H_P}{\left(1+\varphi\right)^{1+q}} \left( G_t \left( 1 + \varphi \left( \omega_t n_t \right)^{\frac{1}{\mu}} \right)^{\mu-1} + (1 - G_t) \right) \left( \varphi + \left( \omega_t n_t \right)^{1+q} \right)^{\psi(1+q)-1},
\]

which replaces (56) and is a rewriting of (163) and \( j_t \) is given by

\[
\hat{j}_t = \left( \frac{q}{1+q} \right)^{\frac{1}{1+q}} \beta G_t \left( 1 + \varphi \left( \omega_t n_t \right)^{\frac{1}{\mu}} \right)^{\mu-1} + (1 - G_t) \frac{H_P}{\left(1+\varphi\right)^{1+q}} \frac{1}{(1+q)^{\frac{1}{1+q}}},
\]

which is derived using (160) and (162).

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

### 7.11 Machines as a capital stock

#### 7.11.1 Set-up

To avoid repetitions, we already include the taxes of section 4.4, namely, we assume that there is a tax \( \tau_m \) on the rental rate of equipment and a tax \( \tau_a \) on high-skill workers in automation innovations. The solution follows similar steps to the baseline case. We denote by \( \tilde{r}_t \) the gross rental rate of machines and by \( \Delta \) their depreciation rate, such that:

\[
\tilde{r}_t = r_t + \Delta. \tag{165}
\]

The Euler equation (24) still applies and the capital accumulation equation is given by (32). The unit cost of product \( i \) is now given by

\[
c(w_L, w_H, \tilde{r}, \alpha(i)) = \left( \frac{w_L^{1-\epsilon} + \alpha(i)\varphi((1+\tau_m)\tilde{r})^{1-\beta_4}w_H^{\beta_4}w_H^{\beta_2}\gamma^{\beta_1\beta_2\beta_3}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}} \right)^{\beta_1} w_H^{\beta_2}\gamma^{\beta_3}, \tag{166}
\]

instead of (4) where \( \varphi \equiv \tilde{\varphi}^{\epsilon} \left( \beta_4^{\beta_1} (1-\beta_4)^{1-\beta_4} \right)^{\epsilon-1}. \) Define \( \mu \equiv \beta_1 (\sigma - 1) / (\epsilon - 1), \) we can then derive the isocost curve as:

\[
N^{1-\sigma} \frac{\sigma}{\sigma - 1} \frac{w_H^{\beta_2}w_H^{\gamma^{\beta_3}}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}} \left( G \left( \varphi \left( ((1+\tau_m)\tilde{r})^{1-\beta_4}w_H^{\beta_4}w_H^{\beta_1} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right) \right) \mu + (1 - G) w_L^{\beta_1(1-\sigma)} \right)^{1-\sigma} = 1. \tag{167}
\]

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

\[
\frac{w_H H^P}{w_L L} = G \left( \beta_2 + \frac{\beta_1 \beta_4 \varphi \left( \left( (1+\tau_m)\tilde{r} \right)^{1-\beta_4}w_H^{\beta_4}w_H^{\beta_1} \right)^{1-\epsilon}}{w_L^{1-\epsilon} + \varphi \left( \left( (1+\tau_m)\tilde{r} \right)^{1-\beta_4}w_H^{\beta_4}w_H^{\beta_1} \right)^{1-\epsilon}} \right) \left( \varphi \left( \left( (1+\tau_m)\tilde{r} \right)^{1-\beta_4}w_H^{\beta_4}w_H^{\beta_1} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right) \mu + \beta_2 (1 - G) w_L^{\beta_1(1-\sigma)} \right) \beta_1 \left( G w_L^{1-\epsilon} \left( \varphi \left( \left( (1+\tau_m)\tilde{r} \right)^{1-\beta_4}w_H^{\beta_4}w_H^{\beta_1} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right) \mu^{-1} + (1 - G) w_L^{\beta_1(1-\sigma)} \right). \tag{168}
\]

118
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:

\[
\frac{\tilde{r}K}{w_H H^P} = \beta_3 + \frac{\beta_1 (1-\beta_4) \varphi \left( \left( \frac{w_L}{w_H} \right)^{1-\beta_4} \right)}{(1+\tau_m) \left( \frac{w_L}{w_H} \right)^{1-\beta_4}} \left( \varphi \left( \left( 1+\tau_m \right) \tilde{r} \right)^{1-\beta_4} \left( \frac{w_H}{w_L} \right)^{1-\epsilon} + \left( \frac{w_H}{w_L} \right)^{1-\epsilon} \right) \left( G \left( \Phi + 1 \right) \beta_4 \right) + \beta_3 (1-G) w_L^{\beta_1 (1-\sigma)}.
\]  

(169)

\[
\frac{w_H H^P}{w_L L} = \frac{G \left( \beta_2 + \beta_1 \beta_4 \left( \frac{\Phi}{\Phi+1} \right) + \beta_2 \left( 1-G \right) \right)}{\beta_1 \left( G \left( \Phi + 1 \right)^{\mu-1} + (1-G) \right)}, \tag{170}
\]

where we defined

\[
\Phi = \varphi \left( \left( \frac{w_L}{w_H} \right)^{1-\beta_4} \right)^{\epsilon-1} = \varphi \left( \left( \frac{w_L}{w_H} \right)^{1-\beta_4} \left( \frac{w_L}{w_H} \beta_4 \right)^{\epsilon-1} \right). \tag{171}
\]

In (170), the RHS is increasing in \(w_L\) and decreasing in \(w_H/w_L\) for given \(G, \tilde{r}\). Therefore, this equation defines the relative demand curve in the \(w_L, w_H\) space as rotating counterclockwise (when \(G > 0\)) when \(w_L\) increases. Plugging (170) in (167) then defines \(w_L\) uniquely as a function of \(N, G, \tilde{r}\) and \(H^P\). We can then derive the effect of changes in \(G\) and \(N\) for given \(H^P\) and \(\tilde{r}\) (i.e. when \(K\) is perfectly elastically supplied) on wages, the skill premium and the labor share as follows:

**Proposition 11.** Consider the equilibrium \((w_L, w_H)\) determined by equations (170) and (167). 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 (167) as:

\[
N^{1-\sigma} \sigma \left( \frac{w_H}{w_L} \right)^{\beta_3} \left( \frac{w_H}{w_L} \right)^{\beta_2} \left( G \left( \varphi \left( \left( \left( 1 + \tau_m \right) \tilde{r} \right)^{1-\beta_4} w_H^{\beta_4} \right) \right) \right)^{1-\epsilon} + w_L^{1-\epsilon} \right)^{\mu} \frac{1}{1-\sigma} = 1.
\] (172)

Using that (170) establishes \( \frac{w_H}{w_L} \) as an increasing function of \( w_L \) otherwise independent of \( N \), we get that (172) implies that \( w_L \) and therefore \( w_H/w_L \) (when \( G > 0 \)) and \( w_H \) itself must increase in \( N \).

(170) 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 (167) the direct effect of an increase in \( G \) is to decrease the LHS (because \( \varphi \left( \left( \left( 1 + \tau_m \right) \tilde{r} \right)^{1-\beta_4} w_H^{\beta_4} \right) \right) \) 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 \( \varphi \left( \left( 1 + \tau_m \right) \tilde{r} \right)^{1-\epsilon} \) replacing \( \varphi \)), 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

\[
Y = \frac{\sigma}{\sigma - 1} \left( w_L L + w_H H^P + \tilde{r} K + T_m \right),
\] (173)
where $T_m$ denotes the tax proceeds from the tax on equipment. We have

$$\frac{T_m}{w_H H^P} = G \frac{\tau_m \beta_1 (1-\beta_4) \varphi \left( ((1+\tau_m) \tilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon}}{(1+\tau_m) \left( w_L^{1-\epsilon} + \varphi \left( ((1+\tau_m) \tilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} \right)} \left( \varphi \left( ((1+\tau_m) \tilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right)^{\mu}.$$  \hfill (174)

Then, we obtain:

$$LS = \frac{\sigma}{\sigma - 1} \left( \frac{w_L L + w_H H^P + \tilde{r} K + T_m}{w_L L + w_H H^P} \right).$$ \hfill (175)

Assume that $H = H^P$, then we get that

$$LS = \frac{\sigma - 1}{\sigma} \left( 1 + \frac{\tilde{r} K + T_m}{w_L L + w_H H^P} \right)^{-1}.$$ \hfill (176)

Using (168), (169), (174) and (171), we obtain:

$$\frac{\tilde{r} K + T_m}{w_L L + w_H H^P} = \frac{G \left( \beta_3 + \beta_1 \Phi \frac{1-\beta_4}{\Phi + 1} \right) \left( \Phi + 1 \right)^{\mu} + \beta_3 \left( 1 - G \right)}{G \left( \beta_2 + \beta_1 \beta_4 + \beta_1 \frac{1-\beta_4}{\Phi + 1} \right) \left( \Phi + 1 \right)^{\mu} + \left( \beta_1 + \beta_2 \right) \left( 1 - G \right)}.$$ \hfill (177)

$$= \frac{\beta_3}{\beta_2 + \beta_1} + \frac{1}{\beta_2 + \beta_1} G \left( \beta_2 + \beta_1 \beta_4 + \beta_1 \frac{1-\beta_4}{\Phi + 1} \right) \left( \beta_1 + \beta_2 \right) \left( 1 - G \right) \left( \Phi + 1 \right)^{-\mu} \frac{1}{1 + \Phi}. \hfill (178)$$

This expression is increasing in $\Phi$. From (170), $\Phi$ 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 $\beta_4$ is close to 1).

Further, we can rewrite (170) as:

$$\frac{w_H H^P}{w_L L} = \frac{\beta_2}{\beta_1} + \frac{\left( \beta_2 + \beta_1 \beta_4 \right)}{\beta_1} \frac{G \Phi \left( \Phi + 1 \right)^{\mu-1}}{G \left( \Phi + 1 \right)^{\mu-1} + \left( 1 - G \right)}.$$ \hfill (179)

We have already derived that an increase in $G$ increases $w_H/w_L$, therefore, this expression shows that it will increase $G \Phi \left( \Phi + 1 \right)^{\mu-1} / G \left( \Phi + 1 \right)^{\mu-1} + \left( 1 - G \right)$. We can then rearrange terms in (176) and
\[ \frac{\tilde{r}K + T_m}{w_L L + w_H H} = \frac{\beta_3}{\beta_2 + \beta_1} + (1 - \beta_4) \frac{\beta_1}{\beta_2 + \beta_1} \left( \beta_1 \beta_4 + \beta_2 + (\beta_2 + \beta_1) \frac{G (\Phi + 1)^{\mu-1} + (1 - G)}{G \Phi (1 + \Phi)^{\mu-1}} \right)^{-1}. \]

The right hand side is an increasing function of \( G \Phi (\Phi + 1)^{\mu-1} \), which ensures that the labor share decreases in \( G \) when \( H \approx HP \).

7.11.3 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 \( \tilde{r} \). Since \( r \) is determined by the Euler equation, then on a path where consumption growth is asymptotically constant, then \( \tilde{r} \) is also asymptotically constant (see (165)). 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 12.** Consider four processes \([N_t]_{t=0}^\infty, [G_t]_{t=0}^\infty, [H_t^P]_{t=0}^\infty, [\tilde{r}_t]_{t=0}^\infty\) where \((N_t, G_t, H_t^P, \tilde{r}_t) \in (0, \infty) \times [0, 1] \times (0, H] \times (0, \infty)\) for all \( t \). Assume that \( G_t, g_t^N, H_t^P\) and \( \tilde{r}_t \) all admit positive and finite limits with \( G_\infty \in (0, 1) \). Then the asymptotic growth rate of high-skill wages \( w_H^t \) and output \( Y_t \) are

\[ g_w^{\infty} = g_Y = g_N^Y / \left[ (\sigma - 1) (\beta_2 + \beta_1 \beta_4) \right], \]

and the asymptotic growth rate of low-skill wages is

\[ g_w^{\infty} = \frac{1 + (\sigma - 1) \beta_1 \beta_4}{1 + (\sigma - 1) \beta_1} g_w^{\infty}. \]

**Proof.** For simplicity we assume that the limits \( g_w^{\infty}, g_w^{N}, g_Y^Y \) exist (although we could show that formally as we did in Appendix 12.2.1). Suppose that \( g_w^{\infty} \leq \beta_4 g_w^{\infty} \). Then \( \Phi_t \) must either tend toward a positive constant or toward 0, in either case (170) implies that \( g_w^{\infty} = g_w^{\infty} \), which is a contradiction as \( \beta_4 < 1 \). Hence it must be that \( g_w^{\infty} > \beta_4 g_w^{\infty} \),
which ensures that $\Phi_t \to \infty$. Using this in (167), we obtain:

$$w_{Ht}^{\beta_2 + \beta_1 \beta_4} \sim_{t \to \infty} \frac{\sigma - 1}{\sigma} \frac{\beta_1 \beta_2 \beta_3 \beta_4 (G_{\infty} \varphi)}{(1 + \tau_m)^{1 - \beta_1} \beta_1 N \beta_1 N_{t}^{-1}}.$$

This establishes $g_{wH}^{uH} = g_{wN}^{uN} / [(\sigma - 1) (\beta_2 + \beta_1 \beta_4)]$, from which we can obtain that $g_{\infty}^{Y} = g_{wH}^{uH} = g_{wN}^{uN} / [(\sigma - 1) (\beta_2 + \beta_1 \beta_4)]$ (using that $H_t^P$ admits a positive limit).

Moreover (170) now implies

$$\frac{w_{Ht} H_{\infty}^P}{w_{Lt} L} \sim_{t \to \infty} \frac{G_{\infty} (\beta_2 + \beta_1 \beta_4)}{\beta_1 (1 - G_{\infty})} \varphi H_{\infty}^P \frac{1 + \beta_1 (\sigma - 1) \beta_3 \beta_4}{\beta_1 (1 - \sigma)},$$

which implies (178). Since $\frac{1 + (\sigma - 1) \beta_1 \beta_4}{1 + (\sigma - 1) \beta_1} > \beta_4$, we verify that $g_{\infty}^{wL} > \beta_4 g_{\infty}^{wH}$.

### 7.11.4 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 $\psi \equiv 1 / [(\sigma - 1) (\beta_2 + \beta_1 \beta_4)]$. We then introduce the same normalized variables as in the baseline model: $\tilde{V}_t^A$, $\tilde{V}_t^N$, $\tilde{\pi}_t^A$, $\tilde{\pi}_t^N$, $\tilde{h}_t^A$, $\tilde{c}_t$ and $\tilde{v}_t$. We also introduce $\tilde{Y}_t \equiv Y_t N^{-\psi}$ and $\tilde{K}_t \equiv K_t N^{-\psi}$. Finally we now define

$$n_t \equiv N_t^{-\frac{1 - \beta_4}{1 + \beta_1 (\sigma - 1) \beta_2 + \beta_1 \beta_4}}$$

and

$$\omega_t \equiv \left( N_t^{-\frac{1 - \beta_4}{1 + \beta_1 (\sigma - 1) \beta_2 + \beta_1 \beta_4}} \frac{\beta_1 (\sigma - 1) \beta_4}{w_{Ht}} \right)^{\frac{1 - \beta_4}{1 + \beta_1 (\sigma - 1) \beta_2 + \beta_1 \beta_4}},$$

so that

$$\left( \frac{w_L}{\tilde{V}_t^{1 - \beta_4} \tilde{V}_t^{1 - \beta_4} w_{Ht}^{1 - \beta_4}} \right)^{\beta_1 (1 - \sigma)} = \omega_t n_t.$$

The transitional dynamics can then be expressed as a system of differential equations in $x_t \equiv \left( n_t, G_t, \tilde{K}_t, \tilde{h}_t^A, \tilde{v}_t, \tilde{c}_t \right)$ where the first three variables are state variables and the last three control variables.
Equation (48) still applies, therefore, we get using (166) that

\[ \hat{\pi}_t^A = \left( \frac{\sigma - 1}{\sigma^\sigma} \right) \left( \varphi \left( \left( (1 + \tau_m) \bar{r}_t \right)^{1 - \beta_4} w_H^\beta_3 \right)^{1 - \epsilon} + w_L^{1 - \epsilon} \right) \left( \frac{w_H^{\beta_2 \beta_3}}{\beta_1 \beta_2 \beta_3} \right)^{1 - \sigma} Y_t. \]

We can rewrite this as

\[ \hat{\pi}_t^A = \left( \frac{\sigma - 1}{\sigma^\sigma} \right) \left( \beta_1 \beta_2 \beta_3 \right)^{-1} \left( \left( \bar{r}_t^{(1 - \beta_4) \beta_2 + \beta_3} \right)^{1 - \sigma} \left( \varphi \left( (1 + \tau_m) \left(1 - \beta_4 \right) (1 - \epsilon) + \left( \omega_t n_t \right)^{1 - \sigma} \right) \left( \frac{1 - \epsilon}{\epsilon} \right) \mu \right)^{1 - \sigma} \right) Y_t. \]

We can derive \( \pi_t^N \) similarly and we find

\[ \hat{\pi}_t^N = \omega_t n_t \left( \varphi \left( (1 + \tau_m) \left(1 - \beta_4 \right) (1 - \epsilon) + \left( \omega_t n_t \right)^{1 - \sigma} \right) \left( \frac{1 - \epsilon}{\epsilon} \right) \mu \right)^{1 - \sigma} \hat{\pi}_t^A. \]

(38) is now replaced by

\[ \dot{n}_t = - \frac{1 - \beta_4}{1 + \beta_1 (\sigma - 1) \beta_2 + \beta_1 \beta_4} g_t^N n_t. \]

(39) still applies and so does (40). Because of the automation tax (41) is replaced by

\[ (r_t - (\psi - 1) g_t^N) \hat{V}_t^N = \pi_t^N + \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa} \left( \hat{V}_t^A - \hat{V}_t^N \right) - (1 + \tau_a) \hat{v}_t \hat{h}_t + \hat{V}_t^N \]

and (42) by

\[ \kappa \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa - 1} \left( \hat{V}_t^A - \hat{V}_t^N \right) = (1 + \tau_a) \hat{v}_t. \]

Combining (180), (182), (183) and (23) with equality, we now obtain:

\[ \hat{v}_t = \hat{v}_t \left( \bar{r}_t - \Delta - (\psi - 1) g_t^N - \gamma \omega_t n_t \left( \varphi \left( (1 + \tau_m) \left(1 - \beta_4 \right) (1 - \epsilon) + \left( \omega_t n_t \right)^{1 - \sigma} \right) \left( \frac{1 - \epsilon}{\epsilon} \right) \mu \right) \left( \frac{1 - \epsilon}{\epsilon} \right) \mu \right)^{1 - \sigma} \hat{h}_t^A \right) \hat{V}_t^N. \]
Following the same steps as those used to derive (46), we now obtain:

\[ \dot{\tilde{h}}^A_t = \gamma \frac{\tilde{h}_t^A}{1 - \kappa} \left( \omega t \left( \varphi \left( 1 + \tau_m \right)^{(1 - \beta_3)(1 - \epsilon)} + \left( \omega t \right)^{1/\mu} \right) \right) - \mu \frac{\tilde{\pi}_t^A}{\tilde{t}} + \left( 1 + \tau_a \right) \frac{1 - \kappa \tilde{h}_t^A}{\kappa} \]

(185)

Further, (24) still applies and we can rewrite it as:

\[ \dot{\tilde{c}}_t = \frac{\tilde{c}_t}{\theta} \left( \tilde{r}_t - \left( \rho + \Delta + \theta \psi g_t^N \right) \right). \]

(186)

Finally, we can rewrite (32) as

\[ \dot{\tilde{K}}_t = \tilde{Y}_t - \tilde{c}_t - \left( \Delta + \psi g_t^N \right) \tilde{K}_t \]

(187)

Equations (181), (39), (184), (185), (186) and (187) form a system of differential equations which depend on \( \tilde{Y}_t, \tilde{\pi}_t^A, \tilde{r}_t \) and \( g_t^N \).

(187) implies

\[ \frac{\sigma}{\sigma - 1} \frac{\tilde{\omega}^{\beta_2 + \beta_4 \beta_1} \tilde{r}^{\beta_3 + \beta_1 (1 - \beta_4)}}{\beta_1 \beta_2 \beta_3} \left( G \left( \varphi \left( 1 + \tau_m \right)^{(1 - \beta_4)(1 - \epsilon)} + \left( \omega t \right)^{1/\mu} \right) \right)^{\mu} + \left( 1 - G \right) \omega t \left( \frac{1}{\tilde{r}^{\beta_3 + \beta_1 (1 - \beta_4)}} \right) = 1, \]

so that

\[ \tilde{r} = \left[ \frac{\sigma - 1 \beta_1 \beta_2 \beta_3}{\sigma} \frac{\tilde{\omega}^{\beta_2 + \beta_4 \beta_1} \tilde{r}^{\beta_3 + \beta_1 (1 - \beta_4)}}{G \left( \varphi \left( 1 + \tau_m \right)^{(1 - \beta_4)(1 - \epsilon)} + \left( \omega t \right)^{1/\mu} \right) \right]^{\mu} + \left( 1 - G \right) \omega t, \]

(188)

which defines \( \tilde{r} \) as a function of \( x_t \) and \( \omega t \). (169) can be written as:

\[ H_t^\mu = \frac{\tilde{r} \tilde{K}_t}{\tilde{v}_t} G_t \left( \beta_2 + \frac{\beta_1 \beta_4 \varphi (1 + \tau_m)^{(1 - \beta_4)(1 - \epsilon)}}{\omega t \left( \varphi (1 + \tau_m)^{(1 - \beta_4)(1 - \epsilon)} + \left( \omega t \right)^{1/\mu} \right)} \right) - \beta_2 \left( 1 - G_t \right) \omega t \left( \frac{\tilde{r}^{\beta_3 + \beta_1 (1 - \beta_4)}}{\tilde{r}^{\beta_3 + \beta_1 (1 - \beta_4)}} \right) \]

(189)
which gives, together with (188), $H_P$ as a function of $x_t$ and $\omega_t$. $g_t^N$ still obeys (55), which then defines it as a function of $x_t$ and $\omega_t$.

Combine (173), (168), (169) and (174) to obtain:

$$\frac{Y}{w_{H}H^P} = \frac{\sigma}{\sigma - 1} \left( G \left( \varphi \left( \left(1 + \tau_m \right) \widehat{\tau} \right)^{1-\beta_4} w_{H}^{\beta_4} \right)^{1-\epsilon} + w_{L}^{1-\epsilon} \right)^{\mu} + (1 - G) w_{L}^{\beta_1(1-\sigma)}$$

which we can rewrite as

$$\hat{Y}_t = \frac{\sigma}{\sigma - 1} \left( \frac{G \left( \varphi \left( \left(1 + \tau_m \right) \widehat{\tau} \right)^{1-\beta_4} \left( \omega_t n_t \right)^{\frac{1}{\sigma}} \right)^{\mu} + (1 - G) \omega_t n_t \right) \hat{v}_t H^P_l.$$

This expression, with the previous equations, gives $\hat{Y}_t$ as a function of $x_t$ and $\omega_t$. (179) then ensures that $\hat{\pi}_t^A$ is defined as a function of $x_t$ and $\omega_t$.

Finally, from (168) we obtain:

$$\omega_t = \left[ \frac{\left( \frac{\beta_4}{H_P L} \right)^{1-\beta_4} \left( \frac{H_P L}{\beta_1} \right) G \left( \varphi \left( \left(1 + \tau_m \right) \widehat{\tau} \right)^{1-\beta_4} \left( \omega_t n_t \right)^{\frac{1}{\sigma}} \right)^{\mu - 1} + (1 - G) \right]^{\frac{\beta_1(1-\sigma)}{1+\beta_1(\sigma-1)}}$$

which implicitly defines $\omega_t$ as a function of $x_t$. Hence, together with (188), (189), (55), (190), (179) and (191), the system formed by (181), (39), (184), (185), (186) and (187) describes the dynamic equilibrium. We then obtain

**Proposition 13.** Assume that

$$\rho \left( \frac{(1 + \tau_m)^{\kappa}}{\kappa^{\gamma} (1 - \kappa)^{1-\kappa}} \left( \frac{\rho}{\gamma} \right)^{1-\kappa} + \frac{1}{\gamma} \right) < \psi H$$

is satisfied, then the economy admits a steady-state $\left( n^*, G^*, \hat{R}^*, \hat{h}^A, \hat{v}, \hat{c}^* \right)$ with $n^* = 0$, $G^* \in (0,1)$ and $g^{N*} > 0$. $g^{N*}, G^*$ and $\hat{h}^A$ are independent of $\tau_m$. 

126
Proof. As before, we directly get that in a steady-state with $g_{N^*} > 0$, we must have $n^* = 0$. (191) then implies that $\omega^*$ is a constant defined by

$$\omega^* = \left[ \left( \frac{\hat{v}^*}{\beta_1} \right)^{1-\beta_1} \frac{H^P*}{L} \frac{\varphi^*}{G^* (\beta_2 + \beta_3 \beta_4)} \right]^{\beta_4 (1-\sigma)}.$$

This guarantees that in such a steady-state, $w_{t+1} = \omega^* \frac{1}{\beta_1 (1-\sigma)} \hat{v}^* \beta_3 \beta_4 (1-\beta_4) \varphi^* \mu^* H^P* \beta_2 (1-\sigma) \sigma_0$. This gives

$$\hat{r}^* = \rho + \Delta + \theta \psi g_{N^*}. \quad (193)$$

(189) implies that

$$H^P* = \frac{\hat{r}^* \hat{K}^*}{\hat{v}^*} (\beta_2 + \beta_3 \beta_4) / (\beta_3 + \beta_1 (1-\beta_4)). \quad (194)$$

Then (190) implies that

$$\hat{Y}^* = \frac{\sigma}{(\sigma - 1) (\beta_2 + \beta_3 \beta_4)} \hat{v}^* H^P*. \quad (195)$$

We then get that (179) implies that

$$\frac{\hat{\pi}^A}{\hat{v}^*} = \frac{(\sigma - 1)^{\sigma - 2} (\beta_1 \beta_2 \beta_3) \sigma - 1}{(\beta_3 + \beta_1 (1-\beta_4))} \left( \frac{\hat{r}^* (1-\beta_4) \beta_3 \beta_2 + \beta_3 \beta_4 \beta_1 (1 + \gamma_m) (1-\beta_4) \beta_1 \varphi \mu H^P* \beta_2 (1-\sigma) \sigma_0}{\hat{v}^* \hat{r}^* \beta_2 (1-\sigma) \sigma_0} \right)^{\beta_4 (1-\sigma)}.$$

(188) gives

$$\hat{r}^* = \left[ \frac{\sigma - 1}{\sigma} \frac{\beta_1 \beta_2 \beta_3 \beta_4}{\hat{v}^* \beta_2 + \beta_4 \beta_1 (1 + \gamma_m) (1-\beta_4) \beta_1} \right]^{\beta_4 (1-\sigma)} \left( \frac{1}{\hat{v}^*} \frac{G^* \varphi \mu}{\hat{v}^*} \right)^{\beta_4 (1-\sigma)}.$$

Therefore (196) simplifies into

$$\frac{\hat{\pi}^A}{\hat{v}^*} = \psi \frac{H^P*}{G^*}, \quad (198)$$

just as in the baseline model. Then (184) and (193) together imply that

$$\hat{h}^A* = \frac{\kappa}{\gamma (1 + \tau_0) (1 - \kappa)} \left( \rho + ((\theta - 1) \psi + 1) g_{N^*} \right). \quad (199)$$

This defines $\hat{h}^A*$ as an increasing function of $g_{N^*}$. Further, in steady-state $G^*$ still obeys
and $H^*$ obeys (60), which imply that $G^*$ and $H^*$ also be defined as function of $g^{N^*}$.

(198), (185), (58), (199) then lead to

$$1 - \frac{\kappa \gamma G^* (1 + \tau_a)}{\psi H^*} \frac{(1 + \tau_a) \tilde{h}^{A^*} (1 - \frac{1}{\gamma})}{\kappa H_t G_t^*} = 1,$$

which up to the term $1 + \tau_a$ is the same as (62) 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 (192) is satisfied. As (200), (58), (60) and (199) are independent of $\tau_m$, so are $g^{N^*}$, $\tilde{h}^{A^*}$ (now given by (199)), $G^*$ (given by (58)) and $H^*$ (given by (60)).

We further obtain $\tilde{r}$ through (193), which must be independent of $\tau_m$ as well. We then get $\tilde{v}$ through (197) as

$$\tilde{v} = \left[ \frac{\sigma - 1}{\sigma} \frac{\beta_1 \beta_2 \beta_3}{\beta_1 \beta_2 \beta_3} \frac{(G^* \varphi^{\mu})^{\frac{1}{\gamma - 1}}}{(1 + \tau_m)^{(1 - \beta_4) \beta_1}} \right]^{\frac{1}{\beta_2 + \beta_4 \beta_1}}.$$

We then get $\tilde{K}$ through (194) and $\tilde{c}$ from (187) which, using (195), implies:

$$\tilde{c} = \frac{\sigma}{(\sigma - 1) (\beta_2 + \beta_1 \beta_4)} \tilde{v} H^* - (\Delta + \psi g^{N^*}) \tilde{K}.$$

Further if $\tau_a = \tau_m = 0$, $g^{N^*}$, $G^*$, $\tilde{h}^{A^*}$ are determined by the same equations are in the baseline model except that the definition of $\psi$ has changed. It is then direct that Proposition 6 extends to this case.

### 7.11.5 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 (171), we can rewrite (167) and (169) as:

$$N^\frac{1}{(\sigma - 1)} \frac{\sigma}{\beta_1 \beta_2 \beta_3} \frac{w_L^{\beta_3} w_H^{\beta_1}}{(G (\Phi + 1)^\mu + 1 - G)^{\frac{1}{\sigma - 1}}} = 1,$$

$$\frac{\tilde{r} K}{w_H H^P} = \frac{G \left( \beta_3 + \frac{\beta_1 (1 - \beta_4) \Phi}{(1 + \tau_m) (1 + \Phi)} \right) (\Phi + 1)^\mu + \beta_3 (1 - G)}{G \left( \beta_2 + \frac{\beta_1 \beta_4 \Phi}{1 + \Phi} \right) (\Phi + 1)^\mu + \beta_2 (1 - G)}.$$

128
Then, totally log-differentiate (171), (170) and (202) to get:

\[
\frac{1}{\varepsilon - 1} \Phi = \hat{w}_L - \hat{w}_H + (1 - \beta_4) \left( \hat{w}_H - \hat{r} \right) - (1 - \beta_4) 1 + \tau_m \tag{203}
\]

\[
\hat{w}_H - \hat{w}_L = \left( \frac{G(\beta_2 + \beta_3 \beta_4 \frac{\Phi}{1+\Phi}) (\Phi+1)^\mu}{G(\beta_2 + \beta_3 \beta_4 \frac{\Phi}{1+\Phi}) (\Phi+1)^\mu + \beta_3 (1-G)} \right) \left( \mu + \frac{\beta_1 \beta_4 \frac{1}{1+\Phi}}{\beta_2 + \beta_1 \beta_4 \frac{\Phi}{1+\Phi}} \right) \Phi \hat{\Phi} \frac{\Phi + 1}{\Phi + 1}, \tag{204}
\]

\[
\hat{r} - \hat{w}_H = \left( \frac{G \left( \beta_3 + \frac{\beta_1 (1-\beta_4) \Phi}{(1+\tau_m)(1+\Phi)} \right) (\Phi+1)^\mu + \beta_3 (1-G)}{G \left( \beta_3 + \frac{\beta_1 (1-\beta_4) \Phi}{1+\Phi} \right) (\Phi+1)^\mu + \beta_2 (1-G)} \right) \left( \mu + \frac{\beta_1 \beta_4 \frac{1}{1+\Phi}}{\beta_2 + \beta_1 \beta_4 \frac{\Phi}{1+\Phi}} \right) \Phi \hat{\Phi} \frac{\Phi + 1}{1 + \Phi} \tag{205}
\]

Combine (203), (204) and (205) to get:

\[
- \left( \frac{\frac{1}{\varepsilon - 1} 1 + \Phi}{\Phi + 1} + \frac{\beta_4 G(\beta_2 + \beta_3 \beta_4 \frac{\Phi}{1+\Phi}) (\Phi+1)^\mu}{G(\beta_2 + \beta_3 \beta_4 \frac{\Phi}{1+\Phi}) (\Phi+1)^\mu + \beta_2 (1-G)} \right) \left( \mu + \frac{\beta_1 \beta_4 \frac{1}{1+\Phi}}{\beta_2 + \beta_1 \beta_4 \frac{\Phi}{1+\Phi}} \right) \Phi \hat{\Phi} \frac{\Phi + 1}{1 + \Phi}
\]

\[
+ \frac{(1-\mu) G(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1} + (1-G)} + \frac{(1-\beta_4) G \left( \beta_3 + \frac{\beta_1 (1-\beta_4) \Phi}{(1+\tau_m)(1+\Phi)} \right) (\Phi+1)^\mu + \beta_3 (1-G)}{G \left( \beta_3 + \frac{\beta_1 (1-\beta_4) \Phi}{(1+\tau_m)(1+\Phi)} \right) (\Phi+1)^\mu + \beta_2 (1-G)} \left( \mu + \frac{\beta_1 \beta_4 \frac{1}{1+\Phi}}{\beta_2 + \beta_1 \beta_4 \frac{\Phi}{1+\Phi}} \right) \Phi \hat{\Phi} \frac{\Phi + 1}{1 + \Phi}
\]

\[
= - \frac{(1 - \beta_4) \beta_3}{G \left( \beta_3 + \frac{\beta_1 (1-\beta_4) \Phi}{(1+\tau_m)(1+\Phi)} \right) (\Phi+1)^\mu + \beta_3 (1-G)} 1 + \tau_m. \tag{206}
\]

As $\mu < 1$, the coefficient in front of $\hat{\Phi}$ on the LHS is positive, so that $\Phi$ is decreasing in the machine tax $\tau_m$. Using (204), we also get that the skill premium decreases in the machine tax.

Totally log-differentiating (201), one gets

\[
\hat{w}_L = \frac{\mu}{\sigma - 1} \frac{G(\Phi+1)^\mu}{G(\Phi+1)^\mu + 1 - G} \Phi \hat{\Phi} \frac{\Phi + 1}{\Phi + 1} - (1 - \beta_1) (\hat{w}_H - \hat{w}_L) - \beta_3 \left( \hat{r} - \hat{w}_H \right).
\]

129
Plugging (204) and (205) and (206), we get:

\[
\hat{w}_L = - \left( \frac{(1-\mu)G(\Phi+1)^{\beta_1}(1-\Phi)}{G(\Phi+1)^{\beta_1}(1-\Phi)} \right) \left[ 1 - \beta_1 + \frac{G(\Phi+1)^{\beta_1}}{G(\Phi+1)^{\beta_1}(1-\Phi)(1+\Phi)} \beta_1 \Phi \right] - \frac{\beta_3}{\epsilon - 1} G(\Phi+1)^{\beta_3}(1-\Phi) \left( \frac{\Phi + 1}{\Phi + 1 + \tau_m} \right) \].

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 14.** On impact, a tax on machines reduces the skill premium and a small tax on machines increases low-skill wages.

### 7.12 Quantitative Exercise

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 \(\hat{Y}_i(t)\) 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:

\[
\min_b (\log(\hat{Y}(b)) - \log(Y))^\prime W (\log(\hat{Y}(b)) - \log(Y)),
\]

where \(W\) is a diagonal matrix of weights. In a previous version of the paper (Héamous 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.

7.12.1 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:

\[
\frac{\alpha}{1 + e^{\frac{\mu - t}{s}}} + \beta,
\]

where \((\alpha, \beta, \mu, s)\) 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
We add these two values to get our combined stock of equipment (+software) over GDP.

### 7.13 Comparison with KORV

We show formally the claims made in section 4.2 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:

\[
F = Ak_s^\alpha \left( \mu u^\sigma + (1 - \mu) \left( \lambda k_e^\rho + (1 - \lambda) s^\rho \right)^{\sigma/\rho} \right)^{1-\alpha/\sigma},
\]

(207)

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

\[
g_{\pi t} = (1 - \sigma)(g_u - g_s) + \sigma \lambda (g_{k_{ct}} - g_s).
\]

Consequently, if investment specific technological change accelerates (that is there are relatively more and more innovations of that type such that \(g_{k_{ct}}\) 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.

\[43\text{Here we use their equation (4) } g_{\pi t} = (1 - \sigma)(g_u - g_s) + (\sigma - \rho)\lambda \left( \frac{k_{ct}}{s} \right)^\rho (g_{k_{ct}} - g_s), \text{ leaving out the labor augmenting terms, which are not included in their preferred specification and do not reflect capital-skill complementarity. This does not affect the present point.}\]
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 \( E_t \left( q_t q_{t+1} \right) \) 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 11 in Appendix 7.11. Taking first order condition in (207), we get the rental rate on structures:

\[
R_{st} = \alpha k_{st}^{\alpha-1} \left( \mu s_t^\sigma + \left(1 - \mu\right) \left( \lambda k_{et}^\rho + \left(1 - \lambda\right) u_t^\sigma \right)^{\sigma/\rho} \right)^{1-\alpha}. \tag{209}
\]

KORV assume that the returns on both capital stocks must be equal, that is:

\[
1 - \delta_s + R_{st} = E_t \left( \frac{q_t}{q_{t+1}} \right) (1 - \delta_e) + q_t R_{et}, \tag{210}
\]

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 (207) with respect to

\[
F = Ak_s^\sigma \left( \mu s^\sigma + \left(1 - \mu\right) \left( \lambda k_e^\rho + \left(1 - \lambda\right) u^{\sigma/\rho} \right)^{\sigma/\rho} \right)^{1-\alpha} \tag{208}
\]

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.

---

\[44\] KORV briefly discuss a production function where the nests are inverted so that:

\[
F = Ak_s^\sigma \left( \mu s^\sigma + \left(1 - \mu\right) \left( \lambda k_e^\rho + \left(1 - \lambda\right) u^{\sigma/\rho} \right)^{\sigma/\rho} \right)^{1-\alpha} \tag{208}
\]

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.
k_{et}, and using (209), we get:

\[ R_{et} = (1 - \mu) \lambda \left( \lambda + (1 - \lambda) \frac{s^\rho_L}{k_{et}^\rho} \right)^{\frac{1-\rho}{\rho}} (1 - \alpha) \left( \frac{\alpha}{R_{st}} \right)^{\frac{\alpha}{1-\alpha}} \left( \frac{\mu u^\sigma_t}{(\lambda k_{et}^\rho + (1 - \lambda) s^\rho_t)^{\frac{\sigma}{\rho}}} + (1 - \mu) \right)^{\frac{1-\sigma}{\sigma}}, \]

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

\[ w_{Lt} = (1 - \alpha) k_{et}^\alpha \mu u^\sigma_t^{-1} \left( \mu u^\sigma_t + (1 - \mu) (\lambda k_{et}^\rho + (1 - \lambda) s^\rho_t)^{\frac{\sigma}{\rho}} \right)^{\frac{1-\alpha}{\sigma}}. \quad (211) \]

Combining this with (209) gives

\[ w_{Lt} = (1 - \alpha) \left( \frac{\alpha}{R_{st}} \right)^{\frac{\alpha}{1-\alpha}} \mu u^\sigma_t^{-1} \left( \mu u^\sigma_t + (1 - \mu) (\lambda k_{et}^\rho + (1 - \lambda) s^\rho_t)^{\frac{\sigma}{\rho}} \right)^{\frac{1-\sigma}{\sigma}}. \quad (212) \]

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.\(^{45}\) This is true regardless of the parameters \( \sigma \) and \( \rho \) (and therefore also applies to Eden and Gaggl, 2018).