6 Main Appendix (For Online Publication)

6.1 A combined HO-AR model

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

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

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 \( \tilde{\varphi} \) to 1. We assume that \( b(i) = \exp(Bi) \) for some \( B > 0 \) and \( \varsigma \in [0,1] \). \( \varsigma \) represents the share of technological progress in new varieties 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-skill 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 an intermediate \( i \) to an intermediate \( i' \) with a higher index is purely labor augmenting \( (\varsigma = 0) \). As soon as newer intermediates also feature more productive machines if automated \( (\varsigma > 0) \) then growth will be asymptotically unbalanced. This result is in the spirit of Uzawa’s theorem but differs in so far as it refers to technological progress from one intermediate to another instead of aggregate technological progress (which here features increasing the range of intermediate inputs and automating some of them).

Formally, we establish in Appendix 7.3:

**Proposition 5.** 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: \( g^{\omega_H} > g^{\omega_L} \).

6.2 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 the production of intermediate inputs \([h_t(i), h_t(i), x_t(i)]_{i \in [0, N_t], t = 0}\), a time path of intermediate inputs production \([y_t(i)]_{i \in [0, N_t], t = 0}\), time paths of high-skill workers engaged in automation \([h^A_t(i)]_{i \in [0, N_t], t = 0}\), and in horizontal innovation \([H^D_t]_{i = 0}\), time paths of final good production and consumption levels \([Y_t, C_t]_{t = 0}\) such that factor markets clear ((18) holds) and good market clears ((12) holds).

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

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

\[
\dot{n}_t = -\frac{\beta}{(1 - \beta)(1 + \beta(\sigma - 1))} g_t^N n_t. \tag{31}
\]

Rewriting (22) with \(\dot{h}_t^A\) gives:

\[
\dot{G}_t = \eta G_t^\kappa \left(\dot{h}_t^A\right)^\kappa (1 - G_t) - G_t g_t^N. \tag{32}
\]

Defining normalized profits \(\hat{\pi}_t^A \equiv N_t^{1-\psi} \pi_t^A\) and \(\hat{\pi}_t^N \equiv N_t^{1-\psi} \pi_t^N\) 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 (19) and (20) as

\[
(r_t - (\psi - 1) g_t^N) \dot{\hat{V}}_t^A = \hat{\pi}_t^A + \dot{\hat{V}}_t^A, \tag{33}
\]

\[
(r_t - (\psi - 1) g_t^N) \dot{\hat{V}}_t^N = \hat{\pi}_t^N + \eta G_t^\kappa \left(\dot{h}_t^A\right)^\kappa \left(\hat{V}_t^A - \hat{V}_t^N\right) - \hat{v}_t \dot{\hat{h}}_t + \dot{\hat{V}}_t^N. \tag{34}
\]

Equation (21) can similarly be rewritten as:

\[
\kappa \eta G_t^\kappa \left(\dot{h}_t^A\right)^{\kappa-1} \left(\hat{V}_t^A - \hat{V}_t^N\right) = \hat{v}_t. \tag{35}
\]

Equation (23) with equality implies that and \(\hat{V}_t^N = \hat{v}_t/\gamma\), therefore using (35) into (34),
we get:

\[(r_t - (\psi - 1) g_t^N) \hat{v}_t = \gamma \hat{\pi}_t^N + \gamma \frac{1 - \kappa}{\kappa} \hat{v}_t \hat{h}_t^A + \hat{v}_t. \tag{36}\]

Taking the difference between (33) and (34) and using (35) we obtain:

\[(r_t - (\psi - 1) g_t^N) (\hat{V}_t^A - \hat{V}_t^N) = \hat{\pi}_t^A - \hat{\pi}_t^N - \frac{1 - \kappa}{\kappa} \hat{v}_t \hat{h}_t^A + (\hat{V}_t^A - \hat{V}_t^N). \]

Using again (35) we get,

\[(r_t - (\psi - 1) g_t^N) \left( \hat{V}_t^A - \hat{V}_t^N \right) = \gamma \hat{\pi}_t^N + \gamma 1 - \kappa \hat{v}_t \hat{h}_t^A \cdot \hat{v}_t. \tag{37}\]

Using (36), we can rewrite this expression as

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

Using (32), this leads to:

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

\[+ (1 - \kappa) \hat{h}_t^A - \frac{\kappa}{G_t} \left( \eta G_t^\kappa \left( \hat{h}_t^A \right)^{\kappa} (1 - G_t) - G_t g_t^N \right). \tag{38}\]

From the definition of \(\omega_t\) and \(n_t\), we get that \(w_t^{\beta(1-\sigma)} = \omega_t n_t\), so that

\[\hat{\pi}_t^N = \omega_t n_t \left( \varphi + (\omega_t n_t) \frac{1}{\pi} \right)^{-\mu} \hat{\pi}_t^A \]
We can then reorder terms in (37) and use (38) to obtain:

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

(39)

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

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

Rewriting (24) using the definition of \(\hat{c}_t\), leads to

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

Combining this equation with (36) and (38), and using the definitions of \(\chi_t, \hat{v}_t\) and \(\hat{\pi}_t^A\) leads to

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

(40)

Together equations (31), (32), (39) and (40) 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

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

(41)

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

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

(42)

Rearranging terms in (11) gives

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

(43)

Using (8), one further gets:

\[Y_t = \sigma \psi \hat{v}_t H_t^P N_t^\psi.\]

(44)
Therefore, rewriting (42) with (43) and (44), one gets:

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

(45)

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:

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

Using (43), this expression can be rewritten as:

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

(46)

This together with (44) implies that \( \hat{c}_t \) obeys

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

Combining this equation with the definition of \( \chi_t \) and (43), leads to

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

(47)

Using the definition of \( H_t^D \) and \( \hat{h}_t^A \), one can rewrite (18) for high-skill workers as:

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

(48)

Together (45), (47) and (48) 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 the definitions of \( n_t \) and \( \omega_t \) to obtain an implicit definition of \( \omega_t \):

\[
\omega_t = \left[ \left( \frac{\sigma-1}{\sigma} \beta \right)^{\frac{1}{\alpha-1}} \frac{H_P}{L} \left( G_t \left( \varphi + (\omega_t n_t)^{\frac{1}{\alpha}} \right)^{\mu-1} (\omega_t n_t)^{\frac{1}{\alpha} - \mu} + (1 - G_t) \right) \right]^{\frac{1}{\alpha}}.
\]

Therefore eventually the system of differential equations satisfied by \( n_t, G_t, \hat{h}_t^A, \chi_t \) is defined by (31), (32), (39) and (40), with \( \hat{\pi}_t^A/\hat{v}_t, H_P, g_t^N \) and \( \omega_t \) given by (45), (47), (48) and (49).

### 6.3 Negative growth for low-skill wages

This section presents two examples with negative growth low-skill wages. Further results on the transitional dynamics derived using simulations are presented in Appendix 7.7.

We ensure temporary negative growth in low-skill wages in figure 7 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 \( (V_t^A - V_t^N) / (w_{H_t}/N_t) \) has already significantly risen (Panel B). Yet, Phase 2 is more intense once it gets started, partly because of the sharp increase in the productivity of the automation technology (following the increase in \( G_t \)) and partly because low-skill wages are higher. Intense automation puts downward pressure on low-skill wages. At the same time, horizontal innovation drops considerably, both because new firms are less competitive than their automated counterparts, and because the high demand for high-skill workers in automation innovations increases the cost of inventing a new product. This results in a short-lived decline in low-skill wages. Indeed, the decline in \( w_{L_t} \) (and increase in high-skill wage \( w_{H_t} \)) lowers the incentive to automate (Panel B), which in return reduces automation.

---

\(^{30}\text{We choose this value for } \tilde{\kappa} \text{ instead of 0.5, because in that case there is no horizontal innovation for some time periods (that is (23) holds with a strict inequality). This is not an issue in principle but simulating this case would require a different numerical approach.}\)
Figure 7: Transitional Dynamics with temporary decline in low-skill wages with an automation externality. Note: same as for Figure 3 but with an automation externality of $\tilde{\kappa} = 0.49$.

Importantly, low-skill wages can drop for $\tilde{\kappa} = 0$—albeit for a small parameter set—as shown in Figure 8 where low-skill wages slightly decline for a short time period (our numerical investigation suggests that larger declines in the absence of an automation externality need to be associated with periods where horizontal innovation completely ceases). The associated parameters are given in Table 3.

Table 3: Baseline Parameter Specification

<table>
<thead>
<tr>
<th></th>
<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>$\tilde{\varphi}$</th>
<th>$\rho$</th>
<th>$\tilde{\kappa}$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<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>

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 (note that in the period where low-skill wages decline the share of high-skill workers hired in production increases slightly, which has a positive contribution on low-skill wages’ growth rates).
6.4 An endogenous supply response in the skill distribution

We present here an extension of the baseline model with an endogenous supply response in the skill distribution. Specifically, let there be a unit mass of heterogeneous individuals, indexed by \( j \in [0, 1] \) each endowed with \( lH \) units of low-skill labor and \( \Gamma(j) = \tilde{H}(1+q)j^{1/q} \) units of high-skill labor (the important assumption here is the existence of a fat tail of individuals with low ability). The parameter \( q > 0 \) governs the shape of the ability distribution with \( q \to \infty \) implying equal distribution of skills and \( q < \infty \) implying a ranking of increasing endowments of high-skill on \( [0, \tilde{H}(1+q)/q] \). Proposition 3 can be extended to this case and in fact the steady state values \((G^*, \tilde{h}^*, g^N*, \chi^*)\) are the same as in the model with a fixed high-skill labor supply \( \tilde{H} \). Proposition 2 also applies except that the asymptotic growth rate of low-skill wages is higher (see Appendix 7.10):

\[
g_{wL}^{\infty} = \frac{1}{1 + q + \beta(\sigma - 1)} g_{Y}^{\infty}, \quad (50)
\]

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

Figure 9 shows the transitional dynamics for this model when the common parameters
are the same as in Table 1, \( \overline{H} = 1/3 \) (so that \( G^*, \hat{h}^{A*}, g^N, \chi^* \) are the same as in the baseline model), \( l = 1 \) and \( q = 0.3 \). The figure looks similar to Figure 3, 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.

\[ 3. \text{ The figure looks similar to Figure 3, but the gap in steady-state between the low-skill growth rate and the high-skill growth rate is a bit smaller. In addition Panel B shows that the skill ratio increases from Phase 2 and Panel A shows that the growth rate is lower in Phase 1 as the mass of high-skill workers is lower then.} \]

\[ 6.5 \text{ Parameters identification} \]

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 \( \bar{\kappa} \) 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.
6.5.1 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 2007, 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 (159), we get that the skill premium must obey:

$$\frac{w_{H1963}}{w_{L1963}} \approx \frac{\beta_2}{\beta_1} \frac{L_{1963}}{H_{1963} - \frac{1}{z} g_{N1963}}.$$  

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

$$\frac{\beta_2}{\beta_1} \approx \frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}},$$  

(51)

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 (159), (160), and (165), we get that the labor share in 1963 should obey

$$l_{s1963} \approx \frac{\beta_2}{\beta_1} \frac{H}{H - \frac{1}{z} g_{N1963}} + \frac{\sigma}{\sigma - 1} + \frac{\beta_2}{H - \frac{1}{z} g_{N1963}} \sigma,$$

which simplifies into

$$l_{s1963} \approx \frac{\sigma - 1}{\sigma} (\beta_2 + \beta_1),$$  

(52)

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 (51) and (52) to obtain

$$\beta_1 \approx \frac{1}{\frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}} + 1} \frac{l_{s1963}}{1 - \frac{1}{\sigma}},$$  

(53)
so that $\beta_1$ which 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 (168) and (168), we get that if the economy is close to its asymptotic steady-state in 2007, the growth rate of the skill premium is given by

$$g_{2007}^{sp} \approx \frac{\beta_1 (\sigma - 1) (1 - \beta_4)}{1 + \beta_1 (\sigma - 1)} g_{2007}^{GDP}. \tag{54}$$

Using (159), (160) (165), the labor share now obeys:

$$l_{s2007} \approx H \left[ \frac{\sigma H}{(\sigma - 1) (\beta_2 + \beta_1 \beta_4)} - \left( \frac{\sigma}{(\sigma - 1) (\beta_2 + \beta_1 \beta_4)} - 1 \right) \left( \frac{1}{\gamma g_{1963}^N + H_{1963}^A} \right) \right]^{-1},$$

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

$$l_{s2007} \approx \frac{\sigma - 1}{\sigma} (\beta_2 + \beta_1 \beta_4). \tag{55}$$

Combining (51), (52) and (55) we obtain:

$$\beta_4 \approx 1 - \left( 1 - \frac{l_{s2007}}{l_{s1963}} \right) \left( \frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}} + 1 \right). \tag{56}$$

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 $l_{s2007}$ by the lowest value over 1983-2007 (which is 62.2%) and $l_{s1963}$ by the highest value (68.8%). With $\frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}} = 0.562$, we then obtain $\beta_4 \approx 0.85$. 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 2007 (so that $l_{s2007}$ is higher than the asymptotic value of the labor share).

Using (53), (54) and (56) we obtain:

$$\sigma \approx \frac{1}{l_{s1963} \left[ \frac{g_{2007}^{GDP}}{g_{2007}^{sp}} \left( 1 - \frac{l_{s2007}}{l_{s1963}} \right) - \frac{1}{\frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}} + 1} \right]}.$$

that is given the initial wage bill ratio and the labor shares in 1963 and 1964, which informs 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 \( \frac{g^{GDP}_{2007}}{g^{GDP}_{2007}} \) is associated with a larger \( \sigma \). When using the last 5 years to determine \( \frac{g^{GDP}_{2007}}{g^{GDP}_{2007}} \), we find that \( \sigma \approx 5.53 \), while our estimation procedure leads to \( \sigma = 6.7 \).

Given \( \sigma \), one can then find \( \beta_1 \) using (53), we find \( \beta_1 \approx 0.54 \), below but not too far from the estimated value of 0.62 (this approximation is not too sensitive on \( \sigma \) provided that \( \sigma \) is large enough). Using (51), we then obtain \( \beta_2 \approx 0.3 \) which is higher than the estimated value of 0.18, in line with the fact that (51) 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 (184), (185) and the definition of GDP, we obtain that

\[
\frac{K^*}{GDP^*} = \frac{1}{\tilde{r}^*} \frac{\beta_3 + \beta_1 (1 - \beta_4)}{(\beta_2 + \beta_1 \beta_4)} \left( \frac{H}{H^{\text{eff}}} - 1 \right).
\]

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

\[
\frac{K^*_{eq}}{GDP^*} = \frac{\beta_1 (1 - \beta_4)}{\beta_3 + \beta_1 (1 - \beta_4)} \frac{\beta_3 + \beta_1 (1 - \beta_4)}{\beta_3 + \beta_1 (1 - \beta_4)} \left( \frac{H}{H^{\text{eff}}} - 1 \right).
\]

Therefore, assuming that in 2000 (the last year for which we have data on the equipment to GDP ratio), we are close to the steady-state, and that most high-skill workers are in production, we get

\[
\rho + \Delta + \theta g^{GDP}_{2000} = \left( \frac{\sigma - 1}{\sigma} \right) \frac{\beta_1 (1 - \beta_4)}{K_{eq,2000}} \frac{K_{eq,2000}}{GDP_{2000}} \approx 0.0443
\]

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 \( \frac{K_{eq,2000}}{GDP_{2000}} = 1.5 \) (with the actual estimated values for \( \sigma, \beta_1 \) and \( \beta_4 \) we would still find that \( \rho + \Delta + \theta g^{GDP}_{2000} \approx 0.094 \)). Yet, as explained in footnote 28, this level is somewhat arbitrary.
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 on the parameters by 2 per cent and illustrate the resulting effect on the initial (1963) and the final (2007) value of each of the four empirical paths. Table 4 gives the resulting 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 parameters $\beta_1, \beta_2$ 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, reduces the rental rate of capital, which increases the use of capital and thereby increases the skill-premium. $\beta_2$ has the opposite effect on the skill premium in 2007. 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, $\kappa, \bar{\kappa}, \eta$ also have a large effect on the skill premium in 2007.

The initial labor share depends on $\beta_1$, $\beta_2$ and $\beta_4$, the latter having a much larger effect in 2007 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 2007 $\sigma, \beta_1, \beta_2, \beta_4$ 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 2007, 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 2007, a higher $\beta_4$ increases the cost of automation and thereby reduces the stock of $K_{eq}/GDP$. The 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 $G_{2007}$ such that $K_{eq}/GDP$ in 2007 is lower as well.
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

\[
V(\theta_i, \bar{\theta}_{-i}(\theta_i)),
\]

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:

\[
\frac{V(\theta_i, \bar{\theta}_{-i}(\theta_i)) - V(\bar{\theta}_i, \bar{\theta}_{-i}(\theta_i))}{V(\theta_i, \bar{\theta}_{-i}(\theta_i))} / \left( \frac{\theta_i - \bar{\theta}_i}{\bar{\theta}_i} \right)^2 \approx \frac{1}{2} \frac{1}{V(\bar{\theta}_i, \bar{\theta}_{-i}(\theta_i))} \frac{d^2V(\bar{\theta}_i, \bar{\theta}_{-i}(\theta_i))}{d\theta_i^2} \bar{\theta}_i^2.
\]

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.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Skill premium</th>
<th>Labor share</th>
<th>GDP/labor</th>
<th>( K_{eq}/GDP )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma )</td>
<td>-0.1 0.1</td>
<td>0.1 0.0</td>
<td>-0.6 -1.9</td>
<td>2.1 1.1</td>
</tr>
<tr>
<td>( \epsilon )</td>
<td>-0.1 -0.2</td>
<td>0.0 0.0</td>
<td>0.0 -0.1</td>
<td>-0.2 -0.1</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.4 0.0</td>
<td>0.7 0.6</td>
<td>-0.8 -2.3</td>
<td>8.3 1.8</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>0.1 -0.2</td>
<td>0.0 0.1</td>
<td>0.0 0.6</td>
<td>-0.2 -0.5</td>
</tr>
<tr>
<td>( \tilde{\kappa} )</td>
<td>-0.2 -0.4</td>
<td>0.0 0.1</td>
<td>0.0 -0.2</td>
<td>0.3 -0.4</td>
</tr>
<tr>
<td>( \theta )</td>
<td>0.0 -0.1</td>
<td>0.0 0.0</td>
<td>0.0 -0.2</td>
<td>0.1 -0.2</td>
</tr>
<tr>
<td>( \eta )</td>
<td>0.1 0.4</td>
<td>0.0 0.0</td>
<td>0.0 0.2</td>
<td>-0.1 0.2</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>-0.1 -0.4</td>
<td>0.0 0.0</td>
<td>0.0 -0.3</td>
<td>0.3 -0.2</td>
</tr>
<tr>
<td>( \rho )</td>
<td>-0.1 -0.1</td>
<td>0.0 0.0</td>
<td>0.0 -0.4</td>
<td>0.2 -0.3</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.5 -0.2</td>
<td>0.2 0.3</td>
<td>-0.5 -1.4</td>
<td>1.1 -0.1</td>
</tr>
<tr>
<td>( \Delta )</td>
<td>0.0 -0.1</td>
<td>0.0 0.0</td>
<td>0.0 -0.1</td>
<td>0.0 -0.2</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>-0.7 -1.9</td>
<td>0.1 0.6</td>
<td>-0.1 -1.3</td>
<td>-5.1 -3.7</td>
</tr>
<tr>
<td>( N_{1963} )</td>
<td>0.0 0.1</td>
<td>0.0 0.0</td>
<td>0.2 0.2</td>
<td>-0.2 0.0</td>
</tr>
<tr>
<td>( G_{1963} )</td>
<td>0.1 0.0</td>
<td>0.0 0.0</td>
<td>0.0 0.0</td>
<td>0.5 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)

6.5.3 The precision of parameters
\( \rho, \theta, \eta, \gamma \) all govern the growth rate of the economy and are weakly identified individually. \( \Delta \) is also not well identified because it mostly affects the growth rate of the capital stock which also depends on \( \rho \) and \( \theta \) (equation 57). Given that this parameter is the one estimated outside a common range this is a reassuring finding.

### 6.5.4 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).

### 6.5.5 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 quite well out-of-sample. The parameter estimates are very similar and the predicted path are close. The largest difference is in the skill premium, where the model calibrated over the first 30 years captures the upward trend but underestimates the pace. In addition, the parameters are very similar to those in the full sample calibration of Table 2.

<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>( \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>12.8</td>
<td>2.4</td>
<td>2142.1</td>
<td>8.3</td>
<td>2.0</td>
</tr>
</tbody>
</table>

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>( \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>6.7</td>
<td>4.95</td>
<td>0.62</td>
<td>0.64</td>
<td>0.57</td>
<td>1.03</td>
<td>0.40</td>
<td>0.57</td>
<td>0.038</td>
<td>0.17</td>
<td>0.015</td>
<td>0.75</td>
<td>20.4</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 6: Parameters (only matching the first 30 years)
Figure 10: Predicted and empirical time paths for $\tilde{\kappa} = 0$

Figure 11: Predicted and empirical time paths only matching the first 30 years
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

$$\frac{w_H}{w_L} = \frac{1 - \beta}{\beta} \frac{L}{H^P} \frac{G (1 + \varphi w_L^{\epsilon-1})^\mu + 1 - G}{G (1 + \varphi w_L^{\epsilon-1})^{\mu-1} + 1 - G}. \quad (58)$$

Since $0 < \mu < 1$, (58) 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 (58) 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, (58) 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 $G$. Assumes now that $w_L$ increases in $G$, then the right-hand side of (58) 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:

$$w_L = \frac{\sigma-1}{\sigma} \beta \left( \frac{H^P}{L} \right)^{(1-\beta)} N^{\frac{1}{\sigma-1}} \left( G \left(1 + \varphi w_L^{\epsilon-1}\right)^{\mu-1} + (1 - G) \right)^{1-\beta} \times \left( G \left(1 + \varphi w_L^{\epsilon-1}\right)^{\mu} + (1 - G) \right)^{\frac{1}{\sigma-1}-(1-\beta)}, \quad (59)$$
Log differentiating with respect to $G$ one obtains:

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

(60)

where

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

$Den > 0$ as $\epsilon > 1$, $\mu \in (0, 1)$ and $\frac{\beta \varphi w_L^{(\epsilon)}}{(1 + \varphi w_L^{(\epsilon)})^\mu} G (1 + \varphi w_L^{(\epsilon)})^\mu < 1$. In (60) the scale effect term is positive as $(1 + \varphi w_L^{(\epsilon)})^{\mu - 1} - 1 > 0$. 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 $1 - (1 + \varphi w_L^{(\epsilon)})^{\mu - 1} > 0$ 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 $\frac{1 - (1 + \varphi w_L^{(\epsilon)})^{\mu - 1}}{G (1 + \varphi w_L^{(\epsilon)})^\mu + 1 - G} \frac{1}{\frac{(1 + \varphi w_L^{(\epsilon)})^{\mu - 1}}{G (1 + \varphi w_L^{(\epsilon)})^\mu + 1 - G}}$ is large enough. From (59) we get:

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

\[> \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_P}{L} \right)^{(1 - \beta)} N \frac{1}{\sigma - 1} \left( 1 + \varphi w_L^{(\epsilon)} \right)^{1 - \beta},\]
where the last line uses that $G \in [0, 1]$. We then obtain that

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

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

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

we get that

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

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 (59) with respect to both $N$ and $G$ and obtain:

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

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

$$\hat{w}_L^{NT} = \left[ \left( \frac{1-(1+\varphi w_L^{\ell-1})^{\mu-1}}{G(1+\varphi w_L^{\ell-1})^{\mu-1}+1-G} \right) \frac{G \hat{G} + \frac{1}{\sigma-1} \hat{N}}{1-Den} \right] \frac{\hat{N}}{Den}. \tag{61}$$

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

\[
\hat{w}_H - \hat{w}_L = \left( \frac{\mu G (1 + \varphi w^{e-1}_L)^\mu}{G (1 + \varphi w^{e-1}_L)^\mu + 1 - G} + \frac{(1 - \mu) G (1 + \varphi w^{e-1}_L)^{\mu - 1}}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + 1 - G} \right) \frac{(\epsilon - 1) \varphi w^{e-1}_L}{1 + \varphi w^{e-1}_L} \hat{w}_L
\]

\[
+ \left( \frac{G ((1 + \varphi w^{e-1}_L)^\mu - 1)}{G (1 + \varphi w^{e-1}_L)^\mu + 1 - G} + \frac{G (1 - (1 + \varphi w^{e-1}_L)^{\mu - 1})}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + 1 - G} \right) \hat{G}.
\]

Using (61) and that \(\hat{G} = -\hat{N}\), then we get that following a change in the mass of non-automated products (keeping the mass of automated products constant):

\[
\hat{w}^{NT}_H = \left( \left( \frac{1 - \mu) G (1 + \varphi w^{e-1}_L)^{\mu - 1}}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + (1 - G)} \right) \frac{(\epsilon - 1) \varphi w^{e-1}_L}{1 + \varphi w^{e-1}_L} + 1 \right) \hat{N}
\]

\[
\times \left[ \frac{G ((1 - \beta)(1 + \varphi w^{e-1}_L)^{\mu - 1})}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + (1 - G)} + \frac{(1 - \beta)(1 - (1 + \varphi w^{e-1}_L)^{\mu - 1})}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + (1 - G)} \right] \frac{\hat{N}}{\text{Den}}
\]

\[
- \left( \frac{G ((1 + \varphi w^{e-1}_L)^\mu - 1)}{G (1 + \varphi w^{e-1}_L)^\mu + (1 - G)} + \frac{G (1 - (1 + \varphi w^{e-1}_L)^{\mu - 1})}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + (1 - G)} \right) \hat{N}.
\]

We then obtain:

\[
\hat{w}^{NT}_H = \left( \frac{\mu G (1 + \varphi w^{e-1}_L)^\mu}{G (1 + \varphi w^{e-1}_L)^\mu + 1 - G} + \frac{(1 - \mu) G (1 + \varphi w^{e-1}_L)^{\mu - 1}}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + 1 - G} \right) \frac{(\epsilon - 1) \varphi w^{e-1}_L}{1 + \varphi w^{e-1}_L} \hat{w}_L
\]

\[
+ \beta \left( \frac{G ((1 + \varphi w^{e-1}_L)^\mu - 1)}{G (1 + \varphi w^{e-1}_L)^\mu + 1 - G} + \frac{G (1 - (1 + \varphi w^{e-1}_L)^{\mu - 1})}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + 1 - G} \right) \frac{G (1 + \varphi w^{e-1}_L)^{\mu - 1}}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + (1 - G)} \frac{\hat{N}}{\text{Den}} - 1
\]

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

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

\[
\hat{w}^{NT}_H - \hat{w}^{NT}_L = \left[ \frac{(\sigma - 1)^{-1}}{G (1 + \varphi w^{e-1}_L)^\mu + 1 - G} \left( \frac{(1 - \mu)(1 + \varphi w^{e-1}_L)^{\mu - 1}}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + 1 - G} \frac{(\epsilon - 1) \varphi w^{e-1}_L}{1 + \varphi w^{e-1}_L} \right) \right] \frac{G \hat{N}}{\text{Den}}
\]

\[
- (1 - \beta) \left[ \frac{(1 + \varphi w^{e-1}_L)^{\mu - 1}}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + 1 - G} + \frac{1 - (1 + \varphi w^{e-1}_L)^{\mu - 1}}{G (1 + \varphi w^{e-1}_L)^{\mu - 1} + 1 - G} \right] \frac{\hat{N}}{\text{Den}}
\]

59
\[
\hat{w}^N - \hat{w}^N_L = \left[ \left( \frac{\epsilon - 1}{\sigma - 1} - \beta \right) \frac{1}{1 + \varphi w^{-1}_L} - (1 - \beta) \right] \left( 1 + \varphi w^{-1}_L \right)^{\mu - 1} \varphi w^{-1}_L G \hat{N} \frac{1}{(G + \varphi w^{-1}_L)^{\mu + 1} - 1 - G} \text{ Den.}
\]

Therefore an increase in the mass of non-automated products reduces the skill premium (and increases the labor share) if and only if \( 1 - \beta > \left( \frac{\epsilon - 1}{\sigma - 1} - \beta \right) \frac{1}{1 + \varphi w^{-1}_L} \). 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) \( w_L < \varphi^{-1} \): automated firms only use low-skill workers and low-skill wages are given by

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

with a skill premium obeying \( \frac{w_H}{w_L} = \frac{1 - \beta}{\beta} \frac{L}{H_P} \).

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

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

Case iii) \( w_L > \varphi^{-1} \) and all automated firms use machines only, in that case, we get that (59) is replaced by

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

and the skill premium obeys:

\[
\frac{w_H}{w_L} = \frac{1 - \beta}{\beta} \frac{L}{H_P} G \left( w_L \varphi \right)^{\beta(\sigma - 1)} + 1 - G.
\]
One can rewrite (63) as

\[ w_L^{1-\beta} = \frac{\sigma - 1}{\sigma} \left( \frac{H^P}{L} \right)^{(1-\beta)} \frac{N^{\frac{1}{\sigma-1}} (1 - G)^{1-\beta} \left( G + (1 - G) (\tilde{\phi} w_L)^{-\beta (\sigma - 1)} \right)^{\frac{1}{\sigma - 1}} \tilde{\varphi}^\beta}{\left( G (\tilde{\phi} w_L)^{\beta (\sigma - 1)} + 1 - G \right)^{1-\beta}}. \]

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

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

Hence \( w_L \) and \( w_H \) are defined uniquely as functions of \( N, G \) and \( H^P \). If \( N^{\frac{1}{\sigma-1}} < \frac{\sigma}{(\sigma - 1) \beta \tilde{\varphi}} \left( \frac{L}{H^P} \right)^{(1-\beta)} \), we are in case i), if \( N^{\frac{1}{\sigma-1}} (1 - G)^{1-\beta} \leq \frac{\sigma}{(\sigma - 1) \beta \tilde{\varphi}} \left( \frac{L}{H^P} \right)^{(1-\beta)} \leq N^{\frac{1}{\sigma-1}} \) then we are in case ii) and if \( N^{\frac{1}{\sigma-1}} (1 - G)^{1-\beta} > \frac{\sigma}{(\sigma - 1) \beta \tilde{\varphi}} \left( \frac{L}{H^P} \right)^{(1-\beta)} \), 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, (63) 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 (63), one gets:

\[ \tilde{w}_L = \left[ \left( \frac{1}{\sigma - 1} - (1 - \beta) \right) \frac{((\tilde{\phi} w_L)^{\beta (\sigma - 1)} - 1) G}{G (\tilde{\phi} w_L)^{\beta (\sigma - 1)} + 1 - G} - (1 - \beta) \frac{G}{1 - G} \right] \frac{\tilde{G}}{Den} \]

where

\[ Den \equiv 1 - \beta \frac{G (\tilde{\phi} w_L)^{\beta (\sigma - 1)}}{G (\tilde{\phi} w_L)^{\beta (\sigma - 1)} + 1 - G} + \frac{(1 - \beta) \beta (\sigma - 1) G (\tilde{\phi} w_L)^{\beta (\sigma - 1)}}{G (\tilde{\phi} w_L)^{\beta (\sigma - 1)} + 1 - G}. \]

We have

\[ \left( \frac{1}{\sigma - 1} - (1 - \beta) \right) \frac{((\tilde{\phi} w_L)^{\beta (\sigma - 1)} - 1) G}{G (\tilde{\phi} w_L)^{\beta (\sigma - 1)} + 1 - G} - (1 - \beta) \frac{G}{1 - G} < \frac{1}{\sigma - 1} - \frac{1 - \beta}{1 - G}, \]

61
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 (63) with respect to both $N$ and $G$, one gets:

$$\hat{w}_L = \left[ \left( \frac{1}{\sigma - 1} - (1 - \beta) \right) \frac{(\tilde{\varphi}_w L)^{\beta (\sigma - 1)} - 1}{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)} + 1 - G} - (1 - \beta) \frac{G}{1 - G} \right) \hat{G} + \frac{1}{\sigma - 1} \hat{N} \right] \frac{1}{\text{Den.}}.$$

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:

$$\hat{w}_L^{NT} = \left( 1 - \beta \right) \left( \frac{(\tilde{\varphi}_w L)^{\beta (\sigma - 1)} - 1}{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)} + 1 - G} + \frac{G}{1 - G} \right) + \frac{1}{\sigma - 1} \left( 1 - \frac{(\tilde{\varphi}_w L)^{\beta (\sigma - 1)} - 1}{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)} + 1 - G} \right) \hat{N} \frac{\hat{N}}{\text{Den.}}.$$

Therefore $w_L$ increases.

Log-differentiating (64), we get:

$$\hat{w}_H = \frac{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)} - 1}{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)} + 1 - G} \hat{G} + \frac{G}{1 - G} \hat{G} + \left( 1 + \beta (\sigma - 1) \frac{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)}}{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)} + 1 - G} \right) \hat{w}_L.$$

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

$$\hat{w}_H^{NT} = \frac{1}{\sigma - 1} \left( 1 - \frac{(\tilde{\varphi}_w L)^{\beta (\sigma - 1)} - 1}{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)} + 1 - G} \right) \hat{N} \frac{\hat{N}}{\text{Den.}}.$$

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

$$\hat{w}_H^{NT} - \hat{w}_L^{NT} = - \frac{(1 - \beta) G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)}}{G (\tilde{\varphi}_w L)^{\beta (\sigma - 1)} + 1 - G} \frac{\hat{N}}{\text{Den.}}.$$

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 in-
creases $w_L$, increases $w_H$ and decreases $w_H/w_L$ if $N$ is large enough but $G \neq 1$ (so that we are in case iii)).

7.2 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 $g_t^N$ admits a positive limit. Therefore, $w_{Lt}$ must be bounded below, so that (11) gives

$$g_{\infty} w_H = \psi g_{\infty}.$$

Further, using that $H_P$ 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 $\epsilon < \infty$, and the case where $\epsilon = \infty$.

**Subcase with $\epsilon < \infty$.** We use equation (59) 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 $\lim w_{Lt} = \infty$, (59) implies

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

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 (59) gives:

$$w_{Lt} \sim \left( \left( \frac{\sigma - 1}{\sigma \beta} \right)^{1/\beta} \frac{H_P}{L} \left( \varphi^{(\psi-1)} \right) \right)^{1} N_t^{\psi} \left( \varphi^{(\rho-1)} + (1 - G_t) w_L^{(\epsilon-1)(1-\rho)} \right)^{1/\rho}.$$  \hspace{1cm} (66)

Following the assumption of Part B in Proposition 2, we assume that $\lim (1 - G_t) N_t^{\psi(\rho-1)(1-\mu)}$ exists and is finite. Suppose first that $\limsup (1 - G_t) w_L^{(\epsilon-1)(1-\mu)} = \infty$, then there must exist a sequence of $t$’s, denoted $t_n$ for which:

$$w_{Lt_n} \sim \left( \left( \frac{\sigma - 1}{\sigma \beta} \right)^{1/\beta} \frac{H_P}{L} \left( \varphi^{(\psi-1)} \right) \right)^{1/(1+\rho(\sigma-1))} \left( (1 - G_{t_n}) N_t^{\psi} \right)^{1/(1+\rho(\sigma-1))}.$$
Yet, this implies

$$(1 - G_{t_n}) w_{Lt_n}^{(\epsilon - 1)(1 - \mu)} \sim \left( \left( \frac{\sigma - 1}{\sigma} \right) \frac{1}{1 - \beta} \frac{H^P}{L} (\varphi^{\mu})^{\psi - 1} \right)^{\left(\epsilon - 1\right)\left(1 - \mu\right)} \frac{1}{1 + \beta(\sigma - 1)} \times \left(1 - G_{t_n} \right) N_{t_n}^{\psi(\epsilon - 1)(1 - \mu)}$$

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 (66) implies

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

Therefore we get that \( g^{w_L} = \frac{\psi}{\epsilon} g^N = \frac{1}{\epsilon} g^{Y} \).

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

$$w_{Lt_m} \sim \left( \left( \frac{\sigma - 1}{\sigma} \right) \frac{1}{1 - \beta} \frac{H^P}{L} (\varphi^{\mu})^{\psi - 1} \right)^{\frac{1}{\epsilon}} N_{t_m}^{\psi}.$$

This leads to

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

which is only possible if \( \lim (1 - G_{t} N_{t}^{\psi(\epsilon - 1)(1 - \mu)} > 0 \). We denote such a limit by \( \lambda \). Then (66) leads to

$$\left( w_{Lt}^{N_{t}^{-\psi}} \right) \sim \left( \left( \frac{\sigma - 1}{\sigma} \right) \frac{1}{1 - \beta} \frac{H^P}{L} (\varphi^{\mu})^{\psi - 1} \right)^{\frac{1}{\epsilon}} (1 - G_{t_m} \right) N_{t_m}^{\psi(\epsilon - 1)(1 - \mu)},$$

which defines uniquely the limit of \( w_{Lt}^{N_{t}^{-\psi}} \). We then obtain that \( g^{w_L} = \frac{\psi}{\epsilon} g^N \). This completes the proof of part B).

**Subcase with \( \epsilon = \infty \)**. Low skill wages are now defined as described in Appendix

31Expressions 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.
7.1.2. First consider the case where $G_\infty < 1$, then Part A) immediately follows. Assume now that $G_\infty = 1$ and that $\lim (1 - G_t) \psi_t$ exists and is finite. Note first that (62) implies that $w_{Lt}$ must be bounded weakly above $\bar{\varphi}$ in the long-run. As a result, (63) leads to

$$w_{Lt} \sim \left( \frac{\sigma - 1}{\sigma} \beta \bar{\varphi}^{\beta(1 - \psi)} \right)^{\frac{1}{1-\beta}} \frac{H_0^P}{L} \left( (1 - G_t) \psi \right)^{\frac{1}{1+\beta(\sigma-1)}}$$

if $w_{Lt} > \bar{\varphi}$.

Since $\lim (1 - G_t) \psi_t$ exists and is finite, $w_{Lt}$ also admits a finite limit. In particular, if $\lim (1 - G_t) \psi_t = 0$, then $w_{L_\infty} = \bar{\varphi}$.

Case where $G_\infty = 0$ (Part C). If $\lim G_t = 0$ then (59) implies that for $\epsilon < \infty$:

$$w_{Lt} \sim \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_0^P}{L} \right)^{(1-\beta)} N_t^{\frac{1}{\sigma-1}} \left( G_t \left( 1 + \varphi_{Lt} \epsilon \right) + 1 \right)^{\frac{1}{\sigma-1} - (1-\beta)}.$$

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

$$w_{Lt} \sim \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_0^P}{L} \right)^{(1-\beta)} N_t^{\frac{1}{\sigma-1}} \left( G_t w_{Lt} \varphi^{\beta(\sigma-1)} + 1 \right)^{\frac{1}{\sigma-1} - (1-\beta)}.$$

(68)

Note that if $\epsilon = \infty$, then we must be in case iii) when $G_\infty = 0$ and (63) also directly implies (68) (as $\varphi^\mu = \bar{\varphi}^{\beta(\sigma-1)}$ in that case).

Assume that $\lim_{t \to \infty} G_t N_t^\beta = \lambda$ exists and is finite. Then (68) implies:

$$w_{Lt} N_t^{-\frac{1}{\sigma-1}} \sim \frac{\sigma - 1}{\sigma} \beta \left( \frac{H_0^P}{L} \right)^{(1-\beta)} \left( \lambda \left( w_{Lt} N_t^{-\frac{1}{\sigma-1}} \right) \varphi^\mu + 1 \right)^{\frac{1}{\sigma-1} - (1-\beta)},$$

which implies that $\lim w_{Lt} N_t^{-\frac{1}{\sigma-1}}$ exists and is finite as well. Therefore one gets that $g_{L_\infty}^w = g_{\infty}^N / (\sigma - 1)$. Using (58) 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 $[H_t^P]_{t=0}^\infty$, such that $g_t^N$ and $H_t^P$ admit
strictly positive limits. If i) the probability that a new product starts out non-automated
is bounded below away from zero and ii) the intensity at which non-automated firms are
automated is bounded above and below away from zero, then any limit of $G_t$ must have
$0 < G_\infty < 1$.

Note that $G_t N_t$ is the mass of automated firms and let $\nu_{1,t} > 0$ be the intensity at
which non-automated firms are automated at time $t$ and $0 \leq \nu_{2,t} < 1$ be the fraction
of new products introduced at time $t$ that are initially automated. Then $(G_t' N_t) =
\nu_{1,t}(1 - G_t) N_t + \nu_{2,t} \dot{N}_t$ such that $\dot{G}_t = \nu_{1,t}(1 - G_t) - (G_t - \nu_{2,t}) g_{t}^N$. First assume that
$G_\infty = 1$, then if $\nu_{1,t} < \bar{\nu}_1 < \infty$ and $\nu_{2,t} < \bar{\nu}_2 < 1$, we get that $\dot{G}_t$ must be negative for
sufficiently large $t$, which contradicts the assumption that $G_\infty = 1$. Similarly if $G_\infty = 0$,
then having $\nu_{1,t} > \nu$ 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 Proof of Proposition 5

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
\[ g(i,t) = 1 - \exp\left(-\eta\left(t - \frac{i}{n}\right)\right) \text{ for } t \geq \frac{i}{n}, \] (69)
since a product $i$ must be born at time $t = i/n$ (and it is born non-automated).

The unit cost function of intermediate $i$ can be written as:
\[ c_i(w_L, w_H, \alpha(i)) = \left(\min\left(\frac{w_L}{b(i)^{1-\varsigma}}, \frac{1}{\alpha(i)}\right)\right)^\beta w_H^{1-\beta} \frac{w_L^{1-\beta}}{b(i)^{1-\beta(1-\varsigma)}} \frac{\beta}{\beta(1-\beta)^{1-\beta}}. \]
Therefore an automated firm ($\alpha(i) = 1$) will use machines instead of low-skill labor if
$w_L > b(i)^{1-\alpha}$. 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
\[ R_i(w_L, w_H, \alpha(i)) = ((\sigma - 1)/\sigma)^{1-\sigma} c_i(w_L, w_H, \alpha(i))^{1-\sigma} Y. \]
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

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

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

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

\[
\frac{w_{H_t} H}{w_{L_t} L} = \frac{1 - \beta}{\beta} \left[ 1 + \frac{\beta (\sigma - 1)}{\sigma - 1} \frac{\int_0^{N_t} 1_{i<I_t} g(i,t) b(i) (1 - \beta (1 - \varsigma)) (\sigma - 1) \, di}{\int_0^{N_t} (1 - 1_{i<I_t} g(i,t)) b(i) (\sigma - 1) \, di} \right], \tag{70}
\]

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

Similarly, using the price normalization, we obtain:

\[
\frac{\sigma}{\sigma - 1} \frac{w_{H_t}^{1-\beta}}{\beta^2 (1 - \beta)^{1-\beta}} \left( \int_0^{N_t} \left( (1 - 1_{i<I_t} g(i,t)) w_{L_t}^{\beta(1-\sigma)} b(i) (\sigma - 1) + 1_{i<I_t} g(i,t) b(i) (1 - \beta (1 - \varsigma)) (\sigma - 1) \right) \, di \right)^{\frac{1}{\sigma - 1}} = 1, \tag{71}
\]

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 (69) and the expression for \(b(i)\) we can compute the integral in (70) and obtain:

\[
\frac{w_{H_t} H}{w_{L_t} L} = \frac{1 - \beta}{\beta} \left[ 1 + \frac{\exp(B(1 - \beta (1 - \varsigma)) (\sigma - 1)) nt - 1}{\exp(B(1 - \beta (1 - \varsigma)) (\sigma - 1)) nt - \exp(-\eta t)} \frac{\exp(B nt - \exp(-\eta t))}{(B (\sigma - 1) + \frac{\eta}{n})} \right] .
\]

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

\[
\frac{w_{H_t} H}{w_{L_t} L} \sim_{t \to \infty} \frac{1 - \beta}{\beta} \left[ 1 + \frac{(B (\sigma - 1) + \frac{\eta}{n})}{(B (1 - \beta (1 - \varsigma)) (\sigma - 1)) ((B (1 - \beta (1 - \varsigma)) (\sigma - 1) + \frac{\eta}{n})} \right] . \tag{72}
\]
Similarly (71) gives

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

From this, we obtain:

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

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 (73) implies that

\[
w_{Ht} = O \left( \exp(B(1 - \beta(1 - \varsigma)) n t) \right).
\]

Further (72) 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

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

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

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

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

In contrast if $\zeta = 0$, then we have $g^{w_H}_\infty = g^{w_L}_\infty = Bn$ but this case only applies if there exist constant $\overline{w}_H$ and $\overline{w}_L$ such that

$$
\overline{w}_H = (1 - \beta) \left( \frac{\sigma - 1}{\sigma} \beta^\sigma \right) \frac{1}{\beta} \left( w_L^{\beta(1-\sigma)} B (\sigma - 1) + \frac{n}{n} \frac{B (1 - \beta) (\sigma - 1) (B (1 - \beta) (\sigma - 1) + \frac{n}{n})}{B (\sigma - 1)} \right)^\psi,
$$

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

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 (70) implies

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

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

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

Since $nt \geq I_t$, then we must have that $w_{Ht} = O(w_{Lt}) = O(\exp ((1 - \zeta) BI_t))$. Similarly (71) now implies:

$$
\frac{\sigma w_{Ht}^{1-\beta}}{(\sigma - 1) \beta^\beta (1 - \beta)^{1-\beta}} \sim \left( \frac{w_L^{\beta(1-\sigma)} \exp (B(\sigma-1)nt - B(\sigma-1)I_t - 1)}{B(\sigma-1)} + w_L^{\beta(1-\sigma)} \exp (-\eta t) \exp ((B(\sigma-1) + \frac{n}{n})I_t - 1)}{B(\sigma-1) + \frac{n}{n}} \right)^{-1}.
$$

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

$$
\frac{\sigma w_{Ht}^{1-\beta}}{(\sigma - 1) \beta^\beta (1 - \beta)^{1-\beta}} \sim \exp ((1 - (1 - \zeta) \beta) BI_t) \left( \frac{\exp (B(\sigma-1)(nt-I_t) - 1)}{B(\sigma-1) + \frac{n}{n}} + \frac{\exp (-\eta (nt-I_t))}{B(\sigma-1) + \frac{n}{n}} \right)^{-1}.
$$
The left-hand side is of order \( \exp ((1 - \varsigma - \beta (1 - \varsigma)) BI_t) \) while the the right-hand side is of order at least \( \exp ((1 - (1 - \varsigma) \beta) BI_t) \), therefore this situation is only possible if \( \varsigma = 0 \). 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 \( \exp (BI_t) \).

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

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

(76)

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

(77)

(76) can be rewritten as \( \overline{w}_H = \frac{1-\beta}{\beta} \overline{w}_L s_1(\overline{w}_L) \) with

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

We obtain (after a lot of algebra):

\[
s'_1(\overline{w}_L) = \begin{cases} 
\left( \begin{array}{c}
\frac{n}{n+1} \left( \frac{n}{2} \frac{B(\sigma-1)}{B(1-\beta)(\sigma-1)+\frac{\eta}{n}} \right) + \frac{1}{B(1-\beta)(\sigma-1)} + \frac{\beta(\sigma-1)}{B(1-\beta)(\sigma-1)+\frac{\eta}{n}} \frac{n}{n+1} \overline{w}_L^{\eta L} \overline{w}_L^{1-\sigma} B(\sigma-1) \\
+ \left( \overline{w}_L^{1-\sigma} B(\sigma-1) + \frac{1}{B(1-\beta)(\sigma-1)} \right) \overline{w}_L^{1-\sigma} - \frac{1}{B(1-\beta)(\sigma-1)} + \frac{1}{B(1-\beta)(\sigma-1)+\frac{\eta}{n}} \frac{n}{n+1} \overline{w}_L^{\eta L} \overline{w}_L^{1-\sigma} B(\sigma-1) \\
\end{array} \right) \\
\left( \overline{w}_L^{1-\sigma} B(\sigma-1) + \frac{1}{B(1-\beta)(\sigma-1)+\frac{\eta}{n}} \right)^2 \\
\left( \frac{n}{n+1} \frac{B(\sigma-1)}{B(1-\beta)(\sigma-1)+\frac{\eta}{n}} \right) \frac{1}{B(1-\beta)(\sigma-1)} + \frac{1}{B(1-\beta)(\sigma-1)+\frac{\eta}{n}} \right) \\
\end{cases}
\]

which is positive when \( \overline{w}_L \leq 1 \) (as then \( \overline{w}_L^{1-\sigma} - 1 \geq 0 \) and \( 1 - \overline{w}_L^{\eta L} \geq 0 \)), so that \( s_1 \) is increasing in \( \overline{w}_L \) for \( \overline{w}_L \leq 1 \). Similarly (77) can be rewritten as \( \overline{w}_H = (1 - \beta) \left( \frac{\sigma - 1}{\sigma} \beta^\beta \right)^{1-\beta} \overline{w}_L^\psi \),
We get

\[ s_2'(\bar{w}_L) = -\bar{w}_L^{-1}(1-\beta)^{-1} \left[ \left( \frac{B}{(\sigma - 1)} \right) \frac{\bar{w}_L^{m} B \beta (\sigma - 1)}{(B (\sigma - 1)) + \frac{n}{n}} + \beta \frac{\bar{w}_L^{1-\sigma} - 1}{B} \right], \]

which is negative for \( \bar{w}_L \leq 1 \). Therefore (75) and (76) 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 (74) and (77) trace a positive relationship (continuity is also easily verified), checking the limits when \( w_L \to 0, \infty \), 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 inputs, where we denote the use of intermediate inputs as \( y_1(i) \) and \( y_2(i) \), respectively, with \( i \in [0, N] \). The two versions of intermediate input \( i \) are produced by the same intermediate input
supplier using production technologies that differ only in the weight on high-skill labor:

\[ y_k(i) = \left[ l_k(i)^{\epsilon_1} + \alpha(i)(\bar{v}x_k(i))^{\epsilon_2} \right]^{1/\epsilon_1} h_k(i)^{1-\beta_k}, \]

where a subscript, \( k = 1, 2 \), refers to the sector where the input 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 \( \psi_k = (\sigma - 1)^{-1}(1 - \beta_k)^{-1} \)).

**Proposition 6.** Consider three processes \( [N_t]_{t=0}^{\infty}, [G_t]_{t=0}^{\infty} \) and \( [H^P_t]_{t=0}^{\infty} \) where \( (N_t, G_t, H^P_t) \in (0, \infty) \times [0,1] \times (0,H) \) for all \( t \). Assume that \( G_t, g^N_t \) and \( H^P_t \) all admit strictly positive limits, then:

\[ g^p = -\psi_2 (\beta_2 - \beta_1) g^N \]
\[ g^{GDP} = \left[ \psi_1 + \psi_1 \beta_1 (\beta_2 - \beta_1) \right] g^N \]

(78)

and if \( G_\infty < 1 \) then the asymptotic growth rate of \( w_L \) is \(^{32}\)

\[ g^w_L = \frac{1}{1 + \beta_1(\sigma - 1)} \frac{1 - \beta_2 + \beta_1 (\beta_2 - \beta_1) (1 - \psi_1^{-1})}{1 - \beta_2 + \beta_1 (\beta_2 - \beta_1)} g^{GDP} \]

(79)

Proposition 6 naturally reduces to Proposition 2 for the special case of \( \beta_2 = \beta_1 \). 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.

\(^{32}\)If \( G_t \) tends towards 1 sufficiently fast such that \( \lim_{t \to \infty} (1 - G_t) Y_t^{\psi_2 (1 - \rho_1)^{-1}} \) is finite, then \( g^w_L = \frac{1}{\epsilon} \left( 1 - \frac{(\beta_2 - \beta_1) (\epsilon - 1)}{1 - \beta_2 + \beta_1} \right) g^{GDP} \geq g^p \) 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.
Proof. The analysis follows similar steps as in the baseline model. The cost function (4) now becomes
\[ c_k (\alpha (i)) = \beta_k^{-\beta_k} (1 - \beta_k)^{-(1-\beta_k)} (w_L^{1-\epsilon} + \varphi (p^x)^{1-\epsilon} \alpha (i))^\frac{\beta_k}{\gamma} w_H^{1-\beta_k}, \tag{80} \]
for \( k \in \{1, 2\} \) indexing, respectively, the production of final good and machines. As before aggregating (80) and the price normalization gives a “productivity” condition, which replaces (11).
\[ \left( G \left( \frac{w_L^{1-\epsilon} + \varphi (p^x)^{1-\epsilon}}{p^x} \right)^{\mu_1} + (1 - G) \frac{w_L^{\beta_1(1-\sigma)}}{p^x} \right)^\frac{1}{\sigma} w_H^{1-\beta_1} = \frac{\sigma - 1}{\sigma} \beta_1 (1 - \beta_1)^{1-\beta_1} N \frac{1}{N}, \tag{81} \]
where we generalize the definition of \( \mu_k \): \( \mu_k \equiv \frac{\beta_k(\sigma-1)}{\epsilon-1} \). Following the same methodology for the production of machines, we get
\[ \left( G \left( \frac{w_L^{1-\epsilon} + \varphi (p^x)^{1-\epsilon}}{p^x} \right)^{\mu_2} + (1 - G) \frac{w_L^{\beta_2(1-\sigma)}}{p^x} \right)^\frac{1}{\sigma} w_H^{1-\beta_2} = \frac{\sigma - 1}{\sigma} \beta_2 (1 - \beta_2)^{1-\beta_2} N \frac{1}{N} \frac{p^x}{\beta_2^2 (1-\beta_2)^{1-\beta_2}}. \tag{82} \]
Taking the ratio between these two expressions, we get
\[ \frac{\left( \left( \frac{w_L^{1-\epsilon} + \varphi (p^x)^{1-\epsilon}}{p^x} \right)^{\mu_2} + (1 - G) \left( \frac{w_L^{1-\epsilon} + \varphi (p^x)^{1-\epsilon}}{p^x} \right)^{\beta_2(1-\sigma)} \right)^\frac{1}{\sigma} w_H^{1-\beta_2}}{\left( \left( \frac{w_L^{1-\epsilon} + \varphi (p^x)^{1-\epsilon}}{p^x} \right)^{\mu_1} + (1 - G) \left( \frac{w_L^{1-\epsilon} + \varphi (p^x)^{1-\epsilon}}{p^x} \right)^{\beta_1(1-\sigma)} \right)^\frac{1}{\sigma} w_H^{1-\beta_1}} = \frac{\beta_2^2 (1-\beta_2)^{1-\beta_2} (p^x)^{1-\beta_2+\beta_1}}{\beta_1 (1-\beta_1)^{1-\beta_1}}. \tag{83} \]
The share of revenues accruing to machines in the production of intermediate input \( i \) for the usage-\( k \) (i.e for use in the final sector or the machines sector) is given by
\[ \nu_{k,x} (\alpha (i)) = \frac{\sigma - 1}{\sigma} \alpha (i) \beta_k \frac{\varphi (p^x)^{1-\epsilon}}{w_L^{1-\epsilon} + \varphi (p^x)^{1-\epsilon}}. \tag{84} \]
aggregating over all intermediates inputs 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
\[ p^x X = N G (R_1 (1) \nu_{1,x} (1) + R_2 (1) \nu_{2,x} (1)). \tag{85} \]
The zero profit condition in the machines sector gives
\[ p^x X = N (G R_2 (1) + (1 - G) R_2 (0)). \tag{86} \]
Revenues themselves are given by

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

(87)

so that (7) still holds but separately for revenues occurring from each activity and with \( \mu_k \) replacing \( \mu \). Combining (7), (84), (85) and (86), we get

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

(88)

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 intermediate input \( i \) 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:

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

(89)

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

\[ w_H = N (\nu_{1,h} (GR_1 (1) + (1 - G)R_1 (0)) + \nu_{2,h} (GR_2 (1) + (1 - G)R_2 (0))), \]

(90)
Take the ratio between (89) and (90), and use (7) to obtain:

\[
\frac{w_L L}{w_H H^P} = \left\{ \begin{array}{l}
\beta_1 \left( G \left( 1 + \varphi \left( \frac{w_L}{p_x} \right)^{\epsilon^{-1}} \right)^{-1} + (1 - G) \left( 1 + \varphi \left( \frac{w_L}{p_x} \right)^{\epsilon^{-1}} \right)^{-\mu_1} \right) \\
+ \beta_2 \frac{R_2(1)}{R_1(1)} \left( G \left( 1 + \varphi \left( \frac{w_L}{p_x} \right)^{\epsilon^{-1}} \right)^{-1} + (1 - G) \left( 1 + \varphi \left( \frac{w_L}{p_x} \right)^{\epsilon^{-1}} \right)^{-\mu_2} \right) \\
(1 - \beta_1) \left( G + (1 - G) \left( 1 + \varphi \left( \frac{w_L}{p_x} \right)^{\epsilon^{-1}} \right)^{-\mu_1} \right) \\
+ (1 - \beta_2) \frac{R_2(1)}{R_1(1)} \left( G + (1 - G) \left( 1 + \varphi \left( \frac{w_L}{p_x} \right)^{\epsilon^{-1}} \right)^{-\mu_2} \right) \\
\end{array} \right. 
\]

(91)

Together (81), (83), (88) and (91) determine \(w_L, w_H, p_x\) and \(R(2)/R(1)\) given \(N, G\) and \(H^P\).

**Asymptotic behavior for \(\epsilon < 1\).** As the supply of machines is going up and there is imperfect substitutability in production between machines and low-skill labor, any equilibrium must feature \(w_{L\infty}/p_{x\infty} = \infty\) even if \(w_{L\infty} < \infty\). Applying this to (83), we get

\[
(p_x^t)^{1-\beta_2+\beta_1} \sim \frac{\beta_1^\beta_1 (1 - \beta_1)^{1-\beta_1}}{\beta_2^\beta_2 (1 - \beta_2)^{1-\beta_2}} \varphi^{\frac{\beta_2-\beta_1}{1-\sigma}} w_{Ht}^{\beta_1^{1-\beta_2}}. 
\]

(92)

Further plugging this last relationship in (81), we get:

\[
w_{Ht} \sim \left( \frac{\sigma-1}{\sigma} \right)^{1+\frac{\beta_1}{\beta_2}} \varphi^{\psi_1 \frac{\beta_2-\beta_1}{1-\beta_2}} w_{Ht}^{1-\beta_2} \left( \frac{\beta_2}{\beta_2} \right)^{\beta_1 (1 - \beta_2)} \left( 1 + \varphi \left( \frac{w_L}{p_x} \right)^{\epsilon^{-1}} \right)^{-\mu_2} G_t^{\psi_1 (1+\beta_2)} N_t^{\psi_1 (1+\beta_2)} 
\]

(93)

Hence

\[
g_{wH} = \psi_1 \left( 1 + \frac{\beta_1 (\beta_2 - \beta_1)}{(1 - \beta_2)} \right) g_N. 
\]

(94)

Through (88), the revenues of the machines sector and the final good sector are of the same order, which implies that \(Y, p^x X\) and \(w_H\) grow at the same rate. Therefore

\[
g_{GDP}^\infty = g_Y^\infty = g_{wH}^\infty = \psi_1 \left( 1 + \frac{\beta_1 (\beta_2 - \beta_1)}{(1 - \beta_2)} \right) g_N. 
\]

In fact (88) gives

\[
\frac{R_{2,t}(1)}{R_{1,t}(1)} \sim \frac{\frac{\sigma-1}{\sigma} \beta_1}{1 - \frac{\sigma-1}{\sigma} \beta_2}. 
\]

(95)
Using (92) and (93), one further gets:

\[
Pt^x \sim \frac{\beta_1^{\beta_1} (1 - \beta_1)^{1 - \beta_1}}{(\beta_2^{\beta_2} (1 - \beta_2)^{1 - \beta_2})} \varphi^{\psi_2 \mu_1 \frac{(\beta_1 - \beta_2)}{\sigma_1}} \left( \frac{\sigma}{\sigma - 1} \right)^{\frac{\beta_2 - \beta_1}{1 - \beta_2}} G_t^{-\psi_2 (\beta_2 - \beta_1)} N_t^{-\psi_2 (\beta_2 - \beta_1)},
\]

therefore

\[g_{Lt}^p = -\psi_2 (\beta_2 - \beta_1) g_N^p < 0, \quad \text{(96)}\]

since \(\beta_2 > \beta_1\). Using that \(w_{Lt}/p^x = \infty\) and (95) in (91) leads to:

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

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

The reasoning then follows that of Appendix 7.2.1. If \(G_\infty < 1\), then (97) implies

\[
w_{Lt}^{1 + \beta_1(\sigma - 1)} \sim \left( p^t \right)^{(\sigma - 1)\beta_1} \frac{w_{Hi} H_t^P \beta_1 (1 - G_t)}{\varphi^{\mu_1} G_t L \left( 1 - \beta_1 + (1 - \beta_2) \frac{\sigma - 1}{\beta_1 - \beta_2} \right)}, \quad \text{(98)}
\]

which, together with (94) and (96) gives (79).

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 \(\lim \sup (1 - G_t) \left( \frac{w_{Lt}}{p^t} \right)^{(\epsilon - 1)(1 - \mu_1)} = \infty\), then there must be a sub-sequence where (98) is satisfied, which with (94) and (96) 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 \(\lim (1 - G_t) \left( \frac{w_{Lt}}{p^t} \right)^{(\epsilon - 1)(1 - \mu_1)} = 0\), then (97) gives

\[
w_{Lt}^{\epsilon - 1} \sim \left( p^t \right)^{\epsilon - 1} \frac{w_{Hi} H_t^P \beta_1 \beta_2 \frac{\sigma - 1}{\beta_1 - \beta_2}}{\varphi L \left( 1 - \beta_1 + (1 - \beta_2) \frac{\sigma - 1}{\beta_1 - \beta_2} \right)},
\]

76
which implies with (94) and (96) that:

\[
\frac{g_{\infty}^L}{g_{\infty}^{GDP}} = \frac{1}{\epsilon} \left( 1 - \frac{(\beta_2 - \beta_1)(\epsilon - 1)}{(1 - \beta_2 + \beta_1)} \right) g_{\infty}^{GDP}.
\]

Finally, if \( \lim \sup (1 - G_t) w_{Lt}^{(\epsilon - 1)(1 - \mu)} \) is finite but strictly positive, then as in Appendix 7.2.1, one can show that this requires that \( \lim (1 - G_t) \frac{w_{Lt}^{(\epsilon - 1)(1 - \mu)}}{w_{Lt}^{(\epsilon - 1)}} > 0 \), from which we can derive that (99) also holds in that case. This proves Proposition 6 and the associated footnote in the imperfect substitutes case.

**Perfect substitutes case.** In the perfect substitutes case, (81) becomes:

\[
\left( G \bar{\varphi}^{\beta_1(\sigma - 1)}(p^x)^{\beta_1(1-\sigma)} + (1 - G)w_L^{\beta_1(1-\sigma)} \right) \frac{1}{1 - \sigma} w_H^{1 - \beta_1} = \frac{\sigma - 1}{\sigma} \beta_1 (1 - \beta_1)^{1 - \beta_1} N^{\frac{1}{1 - \sigma}} \text{ for } w_L > \frac{p^x}{\bar{\varphi}} \]

(83) becomes

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

\[
p_x = \frac{\beta_1^{\beta_1}(1 - \beta_1)^{1 - \beta_1}}{\beta_2^{\beta_2}(1 - \beta_2)^{1 - \beta_2}} w_L^{\beta_2 - \beta_1} w_H^{1 - \beta_2} \text{ for } w_L < \frac{p^x}{\bar{\varphi}};
\]

(88) becomes

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

with \( R_2(1) = 0 \) for \( w_L < \frac{p^x}{\bar{\varphi}} \); and (91) becomes

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

(105)
\[
\frac{w_L}{w_H L^P} = \frac{\beta_1}{1 - \beta_1} \quad \text{for } w_L < p^x/\bar{\varphi}.
\] (106)

Together (101), (103) and (106) show that we must have \(w_L t \geq \frac{p^x}{\tilde{\varphi}}\) for \(t\) large enough, which delivers (94) and (96).

Assume that \(G_\infty < 1\), then (105) gives (98) from which we get that (79) is satisfied.

Now consider the case where \(G_\infty = 1\) and \(\lim (1 - G_t) N^t_\psi\) exists and is finite. Then (105) and (104) imply

\[
w_L t \sim (1 - G_t) w_H t \left(\frac{\bar{\varphi} w_L}{p^t} \right)^{\beta_1(1-\sigma)} \left(\beta_1 + \beta_2 \frac{\bar{\varphi} w_L}{p^t} \left(\frac{\bar{\varphi} w_L}{p^t} \right)^{-(\beta_2 - \beta_1)(\sigma-1)}\right) \frac{H_L^P}{L} \quad \text{for } w_L > p^x/\bar{\varphi}.
\]

We can then derive that \(\bar{\varphi} w_L / p^t\) must have a finite (and positive) limit, so that

\[
g_w^\infty = g_p^\infty = -\frac{\beta_2 - \beta_1}{1 - \beta_2 + \beta_1} g_{GDP}^\infty.
\]

This proves Proposition 6 and its associated footnote in the perfect substitutes case. \(\square\)

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

#### 7.5.1 Proof of Proposition 3

We look for a steady state with positive long-run growth for the system defined by (31), (32), (39) and (40) and we denote such a (potential) steady state \(n^*, G^*, \hat{h}^A^*, \chi^*\) (more generally we denote all variables at steady state with a \(^*\)).

Following (31), we immediately get that \(n^* = 0\). Using (32), we get that \(G^*\) obeys:

\[
G^* = \frac{\eta (G^*)^{\bar{\kappa}} \left(\hat{h}^A^*\right)^{\kappa}}{\eta (G^*)^{\bar{\kappa}} \left(\hat{h}^A^*\right)^{\kappa} + g_{N}^N}.
\] (107)

We focus on a solution with \(G^* > 0\) (when \(\bar{\kappa} > 0\), \(G^* = 0\) is also a solution), this equation implies that with \((g_{N}^N)^* > 0\), \(G^* < 1\). Then, recalling that \(\mu \in (0, 1)\), (49), implies that:

---

\(^{33}\)Note that if \(g_{N}^N = 0\), 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 \(G_t = 1\).
\[
\omega^* = \left[ \left( \frac{\sigma-1}{\sigma} \right)^{\frac{1}{1-\beta}} \frac{H^P_*}{L^*} (1 - G^*)(G^* \varphi^*)^{\psi-1} \right]^{\frac{\beta (1-\sigma)}{1+\beta (1-\sigma)}}.
\]

Using (45), (40) implies that in steady state,

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

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 \( G^* > 0 \), (107) combined with (108), defines \( G^* \) uniquely as an increasing function of \( g^{N*} \). (48) also uniquely defines \( H^P* \) as a function of \( g^{N*} \):

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

(45) and (107) allows to rewrite (39) in steady state as:

\[
\frac{\eta \kappa (G^*)^{\kappa-1} (\hat{h}^A*)^\kappa}{1 - \kappa} \psi H^P* = \frac{\gamma}{\kappa} \left( \hat{h}^A* \right)^2 + \eta G^*_t \left( \hat{h}^A* \right)^{\kappa+1}.
\]

Since \( G^* \), \( \hat{h}^A* \) and \( H^P* \) are functions of \( g^{N*} \), one can rewrite (110) as an equation determining \( g^{N*} \). A steady state with positive growth-rate is a solution to

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

with \( g^{N*} > 0 \). Indeed, (47) simply determines \( \chi^* \) as:

\[
\chi^* = \left( \frac{\sigma}{\sigma - 1} \right)^{\frac{1}{1-\theta}} (1 - \theta \beta)^{\theta} \left( H^P* \right)^{\theta}
\]

which achieves the characterization of a steady state for the system of differential equations defined by (31), (32), (39) and (40).

In order to establish the sufficiency of equation (25) for positive growth. Note that as \( g^{N*} \to 0 \), then equations (108), (107) and (109) imply that

\[
f (0) = \frac{\rho}{\psi H} \left( \frac{1}{\eta \kappa} \left( \frac{1}{1-\kappa} \right)^{1-\kappa} \left( \frac{\rho}{\gamma} \right)^{1-\kappa} + \frac{1}{\gamma} \right).
\]
In addition, \( \frac{g^N}{\gamma} + (1 - G^*) \hat{h}^{A*} \) is always greater than \( g^N \), 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, \( f (g^N) > 1 \). Therefore a sufficient condition for the existence of at least one steady state with positive growth and positive \( G^* \) is that \( f (0) < 1 \) (such that \( f (g^N) = 1 \) has a solution), which is equivalent to condition (25).

### 7.5.2 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

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

Note that \( H^{P*} \) is decreasing in \( g^N \) and \( \hat{h}^{A*} \) is increasing in \( g^N \), so a sufficient condition for \( f \) to be increasing in \( g^N \) is that \( G^* \hat{h}^{A*} \) is also increasing in \( g^N \). With \( \bar{\kappa} = 0 \), using (108), (107), we get:

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

Therefore

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

Since \( g^N > 0 \), we get that \( \frac{d (G^* \hat{h}^{A*)}}{dg^N} > 0 \) (so that the steady state is unique) if

\[
   \frac{(1 - \kappa) \gamma}{\eta \kappa^\rho} \rho^{1 - \kappa} < (\theta - 1) \psi + 1. \]

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).
7.5.3 Transitional dynamics and the first phase

Combining (27) and (26), we can write:

\[ N_t h^A_t = \left( \kappa \eta G^N_t \left( \int_t^\infty \exp \left( - \int_t^\tau r_u du \right) \left( \frac{N_t}{w_{H_t}} \left( \pi^A_{\tau} - \pi^N_{\tau} \right) \right) - \frac{1 - \kappa}{\kappa} \frac{N_t}{w_{H_t}} \left( N_t h^A_{\tau} \right) \right) \right)^{1 \over 1 - \kappa}. \]

Using (8) and that aggregate profits \( \Pi_t = N_t \left( G_t \pi^A_t + (1 - G_t) \pi^N_t \right) \) are a share 1/\( \sigma \) of output, we can rewrite this equation as:

\[ \hat{h}^A_t = \left( \kappa \eta G^N_t \left( \int_t^\infty \exp \left( - \int_t^\tau r_u du \right) \left( \psi H_t^P \frac{\pi^A_{\tau} - \pi^N_{\tau}}{G_t \pi^A_t + (1 - G_t) \pi^N_t} \right) - \frac{1 - \kappa}{\kappa} \frac{N_t}{w_{H_t}} \left( N_t \hat{h}^A_{\tau} \right) \right) \right)^{1 \over 1 - \kappa}. \]

Recalling (7), we can write:

\[ \hat{h}^A_t = \left( \kappa \eta G^N_t \left( \int_t^\infty \left( \psi H_t^P \left( \frac{1 + \varphi w^\tau_{L_t} - 1}{G_t(1 + \varphi w^\tau_{L_t})} \right)^{\mu - 1} \exp \left( \int_t^\tau \left( g^N_u - r_u \right) du \right) \right) - \frac{1 - \kappa}{\kappa} \exp \left( \int_t^\tau \left( g^N_u - g^N_u - r_u \right) du \right) \hat{h}^A_{\tau} \right) \right)^{1 \over 1 - \kappa}. \]

Consider a fixed \( \hat{t} > 0 \). Then for an arbitrarily large \( T \), if \( w_{L0} \) is sufficiently small relative to \( \bar{z}^{-1} \), we will have that \( w_{L_t} \) is small relative to \( \bar{z}^{-1} \) over \( (0, \hat{t} + T) \). For any \( \tau \in (0, \hat{t} + T) \), we have that \( \left( \frac{1 + \varphi w^\tau_{L_t} - 1}{G_t(1 + \varphi w^\tau_{L_t})} \right)^{\mu - 1} = \mu \varphi w^\tau_{L_t} + o \left( \varphi w^\tau_{L_t} \right) \). 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 \( \limsup |f(z) / z| < \infty \)). Then for any \( t \in (0, \hat{t}) \)

\[ \left( \hat{h}^A_t \right)^{1 - \kappa} \leq \kappa \eta G^N_t \left( \int_t^{\hat{t} + T} \psi H_t^P \left( \mu \varphi w^\tau_{L_t} + o \left( \varphi w^\tau_{L_t} \right) \right) \exp \left( \int_t^\tau \left( g^N_u - r_u \right) du \right) d\tau \right). \]

Further, we know that \( r_u = \rho + \theta g^C_u \) 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 \( \pi^N_u \) initially grows like \( Y_u / N_u \) (and from then on will grow
slower), therefore we have that \( r_u - \varphi g^N_u > \rho \). Hence one can write:

\[
\left( h_t^A \right)^{1-\kappa} \leq \kappa \eta G_t^{\kappa} \left( \int_t^{\hat{\tau}+T} \mu \psi H^P_t \varphi w^{N+1}_{L+T} \exp \left( \int_t^\tau \left( g^N_u - r_u \right) d\tau \right) d\tau + o \left( \varphi w^{N+1}_{L+T} \right) + o \left( e^{-\rho \left( T+\hat{\tau} - t \right)} \right) \right)
\]

Since \( r_u - \varphi g^N_u > \rho \), there exists a \( \phi > 0 \), such that

\[
\int_t^{\hat{\tau}+T} \exp \left( \int_t^\tau \left( g^N_u - r_u \right) d\tau \right) d\tau \leq \int_t^{\hat{\tau}+T} e^{-\phi \left( \tau-t \right)} d\tau \leq \frac{1}{\rho + \phi} \left( 1 - e^{-\phi \left( \hat{\tau}+T - t \right)} \right).
\]

This allows us to rewrite:

\[
\left( h_t^A \right)^{1-\kappa} \leq \kappa \eta G_t^{\kappa} \left( \frac{\mu \psi H^P_t \varphi w^{N+1}_{L+T}}{\rho} + o \left( \varphi w^{N+1}_{L+T} \right) + o \left( e^{-\rho T} \right) \right)
\]

Therefore, since \( T \) is large and \( \varphi w^{N+1}_{L+T} \) is small, then \( h_t^A \) must be small too. In fact, we get that \( h_t^A = O \left( \left( \varphi w^{N+1}_{L+T} \right)^{1-\kappa} \right) + o \left( e^{-\rho T} \right) \).

For any \( t \in (0, \hat{\tau}) \), we can then rewrite (40) as

\[
\frac{\dot{X}_t}{\chi_t} = \gamma \psi H^P - \rho - \left( \theta \psi - \psi + 1 \right) g_t^N + O \left( \left( \varphi w^{N+1}_{L+T} \right)^{1-\kappa} \right) + o \left( e^{-\rho T} \right).
\]

(115)

Using (46) we obtain

\[
C_t = Y_t - X_t = \left( 1 + O \left( G_t \varphi w^{N+1}_{L+T} \right) \right) Y_t.
\]

Next (5) and the corresponding equation for high-skill labor demand in production imply:

\[
\frac{L^{NA}}{L^A} = \frac{\left( 1 - G_t \right) \left( 1 + \varphi w^{N+1}_{L+T} \right)^{-\mu-1}}{G_t} \quad \text{and} \quad \frac{H^{P,NA}}{H^{P,A}} = \frac{\left( 1 - G_t \right) \left( 1 + \varphi w^{N+1}_{L+T} \right)^{-\mu}}{G_t}.
\]

Using (3), we can then write

\[
Y_t = N^{\frac{1}{\sigma-1}} L^\beta \left( H^P_t \right)^{1-\beta} \times
\]

82
\[
\left( G_t \left[ 1 + O \left( \varphi w_{Lt}^{\epsilon-1} \right) + \left( O \left( \varphi w_{Lt}^{\epsilon-1} \right) \varphi_L Y_t \right)^{\frac{\epsilon-1}{\epsilon}} \right] \right)^{\frac{\sigma}{\sigma-1} \frac{\sigma-1}{\epsilon-1}} (1 + O \left( \varphi w_{Lt}^{\epsilon-1} \right)) + 1 - G + O \left( \varphi w_{Lt}^{\epsilon-1} \right) \]

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

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

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

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

while for low-skill wages, we get

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

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

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

Differentiating and plugging into (115) and using (48), we get (recalling (116) so that \( d\ln (1 + O \left( \varphi w_{Lt}^{\epsilon-1} \right)) / dt \) will be of order \( O \left( \varphi w_{Lt}^{\epsilon-1} \right) \) as well).

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

we dropped terms in \( \varphi w_{Lt}^{\epsilon-1} \) since there will negligible in front of \( \left( \varphi w_{Lt}^{\epsilon-1} \right)^{\frac{1}{1-\kappa}} \). 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 \((0, \hat{t})\)

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

and \( g_t^N = \frac{\gamma H \psi - \rho}{\psi + \frac{\theta-1}{\sigma-1} + 1} + O \left( \left( \varphi w_{Lt}^{\epsilon-1} \right)^{\frac{1}{1-\kappa}} \right) + o (e^{-\rho T}) \)

which is positive under assumption (25). We then have that for \( N_t \) low, (22) can be solved
as $G_t = G_0 \exp \left( -\frac{\gamma H \psi - \rho}{\psi + \frac{\sigma}{2} + 1} t \right) + O \left( \left( \varphi w_{Lt+T}^{\psi-1} \right)^{\frac{1}{\kappa}} \right) + o \left( e^{-\rho T} \right)$. This characterizes the solution during Phase 1. We can summarize our discussion in the following Proposition.

**Proposition 7.** For $N_0$ low enough that $w_{L0}$ is small relative to $\varphi^{1/(\epsilon - 1)}$, then there is an interval $(0, \hat{t})$ during which the automation rate $\eta G_t \left( \hat{h}_t^A \right)^\kappa$ is small, the share of automated products $G_t$ depreciates and the economy behaves similarly to that of a Romer model where automation is impossible.

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

We prove the following Proposition:

**Proposition 8.** If $\kappa (1 - \beta) + \tilde{\kappa} < 1$, then $G_t$ cannot go toward 0.

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

When $\tilde{\kappa} > 0$ (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). We look for a sufficient condition under which the Poisson rate will take off.

First assume that $G_t w_{Lt}^{\beta(\sigma - 1)}$ does not tend towards 0. Then from (114) we obtain that:

$$\hat{h}_t^A = \hat{h}_t^A \left( G_t^{\tilde{\kappa} - 1} \right)^{\frac{1}{\kappa}} \implies \eta G_t^{\tilde{\kappa}} \left( \hat{h}_t^A \right)^\kappa = O \left( G_t^{\frac{\tilde{\kappa} - \kappa}{1 - \kappa}} \right)$$

(119)

Since $\tilde{\kappa} \leq \kappa$, we obtain that the Poisson rate of automation diverges: a contradiction.

Assume now that $G_t w_{Lt}^{\beta(\sigma - 1)}$ does tend towards 0. This ensures that $w_{Lt} = O \left( N_t^{\frac{1}{2 - \kappa}} \right)$. Moreover, $\frac{\pi_t^A - \pi_t^N}{G_t \pi_t^A + (1-G_t) \pi_t^N} = O \left( w_{Lt}^{\beta(\sigma - 1)} \right)$. Then using this in (114), we obtain

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

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

\[ \eta \tilde{G}_t \left( \tilde{h}^A_t \right)^\kappa = O \left( G_t^{1-\kappa} N_t^{\frac{\beta}{1-\kappa}} \right) \]

Plugging this in (32) we get:

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

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

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

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

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

\[ \tilde{G}_t^{\tilde{\kappa}} w^{\beta(\sigma-1)} = O \left( N_t^{\frac{\beta(1-\kappa)(1-\tilde{\kappa})}{1-\kappa}} \right), \]

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

Alternatively, if \( G_t^{1-\kappa} N_t^{\frac{\beta}{1-\kappa}} \) goes to 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

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

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

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

### 7.5.5 Transitional dynamics in the third phase

In this appendix, we prove two results:
Lemma 2. If $G_t$ is bounded above 0 then $\hat{h}_t^A$ is bounded.

Proposition 9. If $G_t$ and $g_t^N$ admit positive limits then the economy converges toward a steady-state as described in Proposition 3.

We then provide details on the behavior of the economy close to the steady-state.

Proof of Lemma 2. Assume that $G_t$ is bounded above 0 (in fact the analysis of Phase 2 in Appendix 7.5.4 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 $H_t^s$), then we must have (following a reasoning similar to that in Appendix 2.4), 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 $\hat{c}_t$ are bounded above and below and $\omega_t$ must be bounded above.

Integrating (19), 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:

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

Rewriting this expression with the normalized variables and using (45), we get:

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

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

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

for some constant $M$. This ensures that $\frac{V_t^A - V_t^N}{w_{Ht}/N_t}$ must remain bounded, and following (21), $\hat{h}_t^A$ must be bounded as well.
Proof of Proposition 9. Assume now that $G_t$ has a limit $G_\infty$ and that $g^N_t$ also has a positive limit $g^N_\infty$. Then (32) implies that $G_\infty < 1$. Following Proposition (2), we then get that $w_{Lt} = O\left(N_t^{1+\beta(\sigma-1)}\right)$ or $\omega_t = O\left(1\right)$. Therefore, we can rewrite the system as (32),

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

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

Knowing that

\[
H_t^P = \left(\frac{\sigma - 1}{\sigma}\right)^{\frac{1}{\tau-1}}(\frac{\frac{1}{\tau-1}}{\beta}) \frac{\beta}{\tau-1} (\frac{1}{\beta}) \beta \frac{\beta}{\tau-1} (\frac{1}{\beta}) \frac{1}{\kappa} \frac{1}{\kappa} (G_t \varphi^\mu) \psi \left(\frac{1}{\beta} - 1\right) + O\left(n_t\right), \quad (120)
\]

\[
\frac{\tilde{\pi}_t^A}{\tilde{v}_t} = \frac{\psi H_t^P}{G_t} + O\left(n_t\right), \quad (121)
\]

and (48). Using that $g^N_t$ and $G_t$ have limits in (32) implies that $\tilde{h}_t^A$ must also have a limit. Using (48), this implies that $H_t^P$ must also have a limit and therefore using (120) that $\chi_t$ must have a limit. In other words, the equilibrium path tends toward the steady-state $\left(\tilde{h}_t^A, G^*, \chi^*\right)$ with $n_t \to 0$ defined in Proposition 3.

Behavior close to the steady-state. In this steady-state, using (108) we get

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

Therefore in the third phase, we obtain that $N_t$ grows at rate $g^N_t = g^{N*} + O\left(1\right)$, that the share of automated product obeys $G_t = G^* + O\left(1\right)$, with the mass of high-skill workers in automation given by $H_t^A = (1 - G^*) \tilde{h}_t^{A*} + o\left(1\right)$ and the mass of high-skill workers in production given by $H_t^P = H^P* + o\left(1\right)$, with $H_t^P$ given by (109). Using (43), we obtain
that wages obey:

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

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

Using (121), the profit made by an automated firm are given by

\[
\pi_t^A = \frac{1}{\sigma} \left( \frac{\sigma - 1}{\sigma} \beta^\psi \right)^{\frac{1}{1 - \beta}} H^P \left( G^* \right)^{\psi - 1} N_t^{\psi - 1} + o \left( N_t^{\psi - 1} \right)
\]
while the profits made by a non-automated firm \( \pi_t^N \) are negligible in front of \( \pi_t^A \). Using (33), the value of an automated firm is then simply given by:

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

\[
r^* = \rho + \theta \psi g^N
\]
is the steady-state interest rate. Following (34) and (35), the normalized value of a non-automated firm obeys:

\[
(r_t - (\psi - 1) g_t^N) \hat{V}_t^N = \pi_t^N + (1 - \kappa) \eta G_t^{\frac{\psi}{1 + \kappa} h_t} \left( \hat{V}_t^A - \hat{V}_t^N \right) + \hat{V}_t^N.
\]
Therefore, one gets that for large \( N_t \),

\[
V_t^N = \frac{(1 - \kappa) \eta G^* \hat{h}_t^{A*}}{r^* - (\psi - 1) g^N + (1 - \kappa) \eta G^* \hat{h}_t^{A*}} V_t^A + o \left( N_t^{\psi - 1} \right)
\]
so that asymptotically all the value of a new firm comes from the profits it makes once automated (obviously we still get that in equilibrium \( V_t^N = w_{Ht}/N_t \)).

**Comparing the growth rate in the number of products in Phase 1 and Phase 3.** Use (122), (124) and (123) to get that in steady-state:

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

$$f = \frac{(1 - \kappa) \eta G^* \hat{h}^A}{\rho + (\psi (\theta - 1) + 1) g_N^* + (1 - \kappa) \eta G^* \hat{h}^A},$$  \hspace{1cm} (125)$$

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

Using $\hat{V}^N = \hat{\psi}/\gamma$ and (121), we then obtain:

$$\frac{f}{\rho + (\psi (\theta - 1) + 1) g_N^*} \frac{\psi H^*}{G^*} = \frac{1}{\gamma}.$$

Rearranging terms, this leads to

$$g_N^* = \frac{f \gamma \psi (H - (1 - G^*) \hat{h}^A) - \rho}{\frac{f}{\gamma} \psi + \frac{\theta - 1}{(\sigma - 1)(1 - \beta)} + 1},$$  \hspace{1cm} (126)$$

while from (118) the growth rate in the first period is approximately given by

$$g_{N1} = \frac{\gamma H \psi - \rho}{\psi + \frac{\theta - 1}{\sigma - 1} + 1}.$$

(127)

The two expressions differ by three terms: In the numerator, $H - (1 - G^*) \hat{h}^A$ in (126) replaces $H$ in (127), 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, $\frac{\theta - 1}{(\sigma - 1)(1 - \beta)}$ in (126) replaces $\frac{\theta - 1}{\sigma - 1}$ in (127), 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 (126) does not exist in (127). 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 and this reduces the growth rate in the number of products, on the other hand a lower $G^*$ reduces the demand for high-skill workers and therefore high-skill wages compared to the profits an automated firm, which increases the asymptotic growth rate.
in the number of products. Combining (32), (107) and (125), we get

\[ \frac{f}{G^*} < 1 \iff \frac{1}{G} (1 - \kappa) \eta G^* \tilde{h}^A^* < \rho + (\psi (\theta - 1) + 1) g_N^* + (1 - \kappa) \eta G^* \tilde{h}^A^* \]

\[ \iff \frac{1 - G}{G} (1 - \kappa) \eta G^* \tilde{h}^A^* < \rho + (\psi (\theta - 1) + 1) g_N^* , \]

\[ \iff (1 - \kappa) g_N^* < \rho + (\psi (\theta - 1) + 1) g_N^* . \]

Since \( \theta \geq 1 \) and \( \kappa < 1 \), this equation necessarily holds. Therefore \( f < G^* \) and it is always the case that \( g_N^* > g_N^1 \).

### 7.5.6 Comparative statics

In this section, we prove Proposition 4. 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 (111), where \( G^*, \tilde{h}^A^* \) 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 \( \gamma \).** (108) implies that \( \tilde{h}^A^* \) is inversely proportional to \( \gamma \) (for given \( g_N^* \)). Formally, we have:

\[ \frac{\partial \tilde{h}^A^*}{\partial \gamma} = -\frac{\tilde{h}^A^*}{\gamma} . \tag{128} \]

Differentiating (107) and using (128) leads to:

\[ \frac{\partial G^*}{\partial \gamma} = -\frac{\kappa g_N^* G^*}{\gamma \left( \eta (G^*) \tilde{h}^A^* \right)^\kappa + (1 - \tilde{\kappa}) g_N^*} , \tag{129} \]

so that for a given \( g_N^* \), \( G^* \) is also decreasing in \( \gamma \). Using (109), (128) and (129), we get:

\[ \frac{\partial H^P^*}{\partial \gamma} = \frac{1}{\gamma} \left( \frac{g_N^*}{\gamma} + (1 - G^*) (1 - \kappa) \tilde{h}^A^* \right. \]

\[ \left. + \frac{(1 - \kappa) \eta G^* \tilde{h}^A^* (g_N^*)^\gamma}{(\eta (G^*)^\kappa (h^A^*)^\kappa + (1 - \tilde{\kappa}) g_N^*)} \right) > 0 \]
so that $H^{P*}$ is increasing in $\gamma$. Note that $f$, defined in (111), can be rewritten as

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

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, (111) implies that $g_N^*$ increases in $\gamma$.

**Comparative static with respect to $\eta$.** For given $g_N^*$, (108) implies that $\hat{h}^{A*}$ does not depend on $\eta$. Differentiating (107), we get:

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

(130)

so for given $g_N^*$, $G^*$ increases in $\eta$. (109) 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 (130) and (111), we obtain:

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

Using (108), we can rewrite this as:

$$\frac{\partial \ln f}{\partial \ln \eta} = \left\{ \begin{array}{c}
\frac{-\eta (G^*)^\kappa \left( \hat{h}^{A*} \right)^\kappa + (1 - \kappa) g_N^*}{\eta (G^*)^\kappa \left( \hat{h}^{A*} \right)^\kappa + (1 - \kappa) g_N^*} \times \left( \frac{g_N^* G^*\hat{h}^{A*}}{H^{P*}} + \frac{\rho + ((\theta - 1)\psi + \kappa) g_N^*}{\gamma (1 - \kappa)} \right) \\
\frac{1}{\eta \lambda G^* G_z} \left( \gamma (1 - \kappa) \left( \hat{h}^{A*} \right)^{1-\kappa} + \frac{1}{\gamma} \right) \end{array} \right\},$$

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 $\eta$. 

91
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 (49) defines \( \omega \) implicitly. We can therefore write equations (31), (32), (39) and (40) as a system of autonomous differential equations

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

with initial conditions on state variables as \((n_0, G_0)\) and an auxiliary equation of \( \omega_t = \dot{\vartheta}(x_t) \).

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

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

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

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

As shown in Appendix 7.5.1 for a set of parameter values, the system admits an asymptotic steady state. We assume that the system has reached this asymptotic steady state by time \( T \) and restrict \( \hat{h}_T^A \) 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:

\[
\hat{x} = \arg\min_x S(x)'WS(x),
\]

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

![Graphs](image.png)

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 3.A), which is in line with a stable ratio of total R&D expenses over GDP across the three phases (Figure 3.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 (fell 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?\textsuperscript{34} In Phase

\textsuperscript{34}More specifically we can write $w_{Lt} = f(N_t, G_t, H^P_t)$, using equations (10) and (11). Differentiating with respect to time and using equation (32) gives:

$$
\frac{\partial^{\gamma D_H} w_{Lt}}{\partial t} = \left( \frac{N_t}{w_{Lt}} \frac{\partial f}{\partial N} - \frac{G_t}{w_{Lt}} \frac{\partial f}{\partial G} \right) + \frac{1}{w_{Lt}} \frac{\partial f}{\partial G} \eta G_t^\alpha (1 - G_t)(\hat{h}_t^A)^\kappa + \frac{1}{w_{Lt}} \frac{\partial f}{\partial H^P} \hat{H}_t^P.
$$

94
1, there is little automation, so wage growth for both skill-groups is driven almost entirely by horizontal innovation. In Phase 2, automation sets in. Low-skill labor is then continuously reallocated from existing products which get automated, to new, not yet automated, products. The immediate impact of automation on low-skill wages is negative, while horizontal innovation has a positive impact, as it both increases the range of available products and decreases the share of automated products. The figure also shows that automation plays an increasing role in explaining the instantaneous growth rate of high-skill wages, while the contribution of horizontal innovation declines. This is because new products capture a smaller and smaller share of the market and therefore do not contribute much to the demand for high-skill labor. Consequently, automation is skill-biased while horizontal innovation is unskilled-biased. We stress that this growth decomposition captures the immediate effect of automation or horizontal innovation. This should not be interpreted as “automation being harmful” to low-skill workers in general. In fact, as we demonstrate in Section 3.7, 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 Decreasing growth rate

Figure 14 shows a case where growth is substantially lower in Phase 3 than in Phase 1 and in fact continuously decreases from the middle of Phase 2. The parameters are identical to the baseline case except for: $\sigma = 2.5$, $\beta = 0.55$, $\eta = 0.1$, $\gamma = 0.23$ and $N_0 = 344.25$, these parameters lower the growth rate of the economy particularly in the asymptotic steady state because automation consumes more resources and is less effective as high-skill workers have a larger factor share in production ($N_0$ is higher so as to shorten Phase 1 in the graph). As argued in the text, this case illustrates that our model does not necessarily predict that intense automation needs to be associated with a boost in economic growth. Conversely, making the automation technology more effective (say by reducing the cost share of high-skill workers, $\beta$) could create the opposite pattern of a low initial growth rate followed by a higher eventual growth rate.

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 $w_{Ht}$. 

95
7.7.3 A delayed decline in the labor share

Empirically, the drop in the labor share is a more recent phenomenon than the increase in the skill premium. In Figure 15, we choose parameters such that this happens. The automation technology is more productive $\eta = 0.4$; the automation technology is less concave $\kappa = 0.9$ (so that a higher incentive to automate is required to get a significant share of high-skill workers in automation innovation); and all other parameters are identical to the baseline case. In this case, more high-skill workers get allocated to automation during Phase 2: as shown in Panel C automation expenditures represent a much larger share of GDP during Phase 2 than they do in the baseline case. The mass of high-skill workers engaged in production declines during Phase 2. This results first in a sharper increase in the skill premium (the skill premium condition moves further to the left). In addition, the drop in the labor share is delayed since innovation spending are part of $GDP$ while capital income is a constant share of output $Y$. The growth rates of low-skill and high-skill wages start diverging significantly from around year 135 and by year 150, the high-skill wage growth rate is 2pp higher than the low-skill wage growth rate, while the total labor share only start declining from around year 150 and in fact increases slightly before.
7.7.4 The effect of the innovation parameters

Figure 16: Deviations from baseline model for more productive horizontal innovation technology ($\gamma$) and more productive automation technology ($\eta$).

Figure 16 shows the impact (relative to the baseline case) of increasing productivity in the automation technology to $\eta = 0.4$ (from 0.2) and the productivity in the horizontal innovation technology to $\gamma = 0.32$ (instead of 0.3). A higher $\eta$ 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.

### 7.7.5 Technological shocks

Figure 17 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.

### 7.7.6 Systematic comparative statics

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

**Figure 18**: Comparative statics with respect to the elasticity of substitution across products ($\sigma$), the elasticity of substitution between machines and low-skill workers in automated firms ($\epsilon$) and the factor share of low-skill workers and machines in production ($\beta$).

Figures 18.A,B,C show that a higher elasticity of substitution across products $\sigma$ 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 18.D,E,F show that the elasticity of substitution between machines and low-skill workers in automated firms, $\epsilon$, plays a limited role (as long as the assumption $\mu < 1$ is kept), a higher elasticity reduces the growth of low-skill wages and increases that of high-skill wages during Phase 2. Figures 18.G,H,I show that a lower factor share in production for high-skill workers (a higher $\beta$) increases the growth rate of the economy (high-skill wages are lower which favors innovation). As a result Phase 2 occurs sooner. In addition, 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 19: Comparative statics with respect to the discount rate ($\rho$), the inverse elasticity of intertemporal substitution ($\theta$) and the productivity of machines ($\tilde{\phi}$)

Figures 19.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 19.D,E,F). Figures 19.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 20: 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 20.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 20.D,E,F. The role of the automation externality has already been discussed in the text, Figures 20.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 21.A,B,C show the impact of the horizontal innovation parameter $\gamma$, which was already discussed in the text. Figures 21.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 22.A,B,C show that a higher initial number of products simply advance the entire evolution of the economy. Figures 22.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: $N_t^A$ (respectively $N_t^N$) denotes the mass of automated (respectively non-automated) firms, $L_t^A$ (respectively $L_t^N$) is the mass of low-skill workers hired in automated (respectively non-automated) firms, and $H_t^{P,A}$ (respectively $H_t^{P,N}$) is the mass of high-skill workers hired in production in automated (respectively non-automated) firms. The social planner problem can then be written as (we write the
Lagrange multipliers next to each constraint):
\[
\max \int_0^\infty e^{-\rho t} \frac{C_1^{1-\theta}}{1-\theta} dt
\]
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
\[
F \equiv \left(\left(\frac{N_t^A}{\sigma}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{N_t^N}{\sigma}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma-1}{\sigma}} \cdot \frac{\sigma-1}{\sigma}.
\]
\[
\bar{w}_t : L_t^A + L_t^N = L,
\]
\[
\bar{v}_t : H_t^{P,A} + H_t^{P,N} + H_t^A + H_t^D = H,
\]
\[
\bar{\zeta}_t : N_t^A = \gamma \left(N_t^A + N_t^N\right) H_t^D - \eta \left(N_t^A\right)^{\kappa} \left(N_t^N + N_t^A\right)^{\kappa-\tilde{\kappa}} \left(H_t^A\right)^{\kappa} \left(N_t^N\right)^{1-\kappa},
\]
\[
\bar{\xi}_t : N_t^N = \eta \left(N_t^A\right)^{\kappa} \left(N_t^N + N_t^A\right)^{\kappa-\tilde{\kappa}} \left(H_t^A\right)^{\kappa} \left(N_t^N\right)^{1-\kappa},
\]

\textbf{Figure 22:} Comparative statics with respect to the initial number of products $N_0$ and the initial share of automated products $G_0$.
$H_t^D \geq 0$.

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

$$\frac{\partial F}{\partial X_t} = 1, \quad (131)$$

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

$$w_{Lt} = \frac{\partial F}{\partial L_t^A} = \frac{\partial F}{\partial L_t^N} \quad \text{and} \quad w_{Ht} = \frac{\partial F}{\partial H_t^{P,A}} = \frac{\partial F}{\partial H_t^{P,N}}, \quad (132)$$

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 in the production of intermediates, while intermediates themselves are priced with a mark-up as they are provided by a monopolist. It is easy to show that for a given $H_t^P$, the optimal provision of machines and allocation of high-skill and low-skill workers across firms can be obtained if the purchase of all intermediate inputs is subsidized by at rate $1/\sigma$ (a lump-sum tax finances the subsidy).

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

$$\rho \tilde{\zeta}_t - \xi_t = \tilde{\lambda}_t \frac{\partial F}{\partial N_t^N} + \tilde{\zeta}_t \gamma H_t^D + \left( \tilde{\xi}_t - \tilde{\zeta}_t \right) \eta \left( H_t^A \right)^\kappa \left( N_t^N \right)^{-\kappa} \times \left( N_t^A \right)^{\kappa-\tilde{\kappa}-1}, \quad (133)$$

$$\rho \tilde{\xi}_t - \tilde{\xi}_t = \tilde{\lambda}_t \frac{\partial F}{\partial N_t^A} + \tilde{\zeta}_t \gamma H_t^D + \left( \tilde{\xi}_t - \tilde{\zeta}_t \right) \eta \left( H_t^A \right)^\kappa \left( N_t^N \right)^{1-\kappa} \left( N_t^A \right)^{\tilde{\kappa}-1} \left( \tilde{\kappa} N_t^N + \kappa N_t^A \right) \left( N_t^N + N_t^A \right)^{\kappa-\tilde{\kappa}-1}. \quad (134)$$

Interestingly, $\frac{\partial F}{\partial N_t^N}$ and $\frac{\partial F}{\partial N_t^A}$ correspond to the profits realized by a non-automated and an automated firm respectively in the equilibrium once the subsidy to the use of inter-
mediates is implemented. Therefore we denote
\[ \pi_t^N = \frac{\partial F_t}{\partial N_t^N} \text{ and } \pi_t^A = \frac{\partial F_t}{\partial N_t^A} \]

Further the (shadow) interest rate is given by
\[ r_t = \rho + \theta \frac{G_t}{c_t} = \rho - \frac{\lambda_t}{\kappa_t} \]
Using that
\[ H_t^A = (1 - G_t) N_t h_t^A \]
we can rewrite (133) and (134) as:
\[ r_t \zeta_t = \pi_t^N + \zeta_t g_t^N + (\xi_t - \zeta_t) \eta (G_t) \kappa N_t^N \left( h_t^A \right)^\kappa ((1 - \kappa) (1 - G_t) + (1 - \kappa) G_t) + \zeta_t, \quad (135) \]
\[ r_t \xi_t = \pi_t^A + \zeta_t g_t^A + (\xi_t - \zeta_t) \eta (G_t) \kappa N_t^N \left( h_t^A \right)^\kappa (1 - G_t) \left( \frac{1 - G_t}{G_t} + \kappa \right) + \zeta_t. \quad (136) \]

These expressions parallel equations (19) and (20) in the paper. The rental social value of a non-automated firm \((r_t \zeta_t)\) consists of the current value of one intermediate (which equals the profits when the optimal subsidy to the use of intermediates inputs 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 the condition that \(H_t^D \geq 0\)):
\[ w_{Ht} \geq \zeta_t \gamma N_t, \quad (137) \]
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 intermediate equals the cost of creating one. The first-order condition with respect to \(H_t^A\) gives:
\[ w_{Ht} = (\xi_t - \zeta_t) \kappa \eta (G_t) \kappa N_t^N \left( h_t^A \right)^\kappa \left( \frac{1 - G_t}{G_t} + \kappa \right)^{\kappa - 1}. \quad (138) \]
This equation is the counterpart of (21) 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 \( (n_t, \xi_t, \hat{h}_t^A, \chi_t) \) which characterizes the solution (when there is positive growth). Equations (31) and (32) still hold, while equations (39) and (40) are replaced with

\[
\dot{\hat{h}}_t^A = \frac{\gamma}{1 - \kappa} \left( \omega_t n_t \left( \varphi + (\omega_t n_t)^{\frac{1}{\beta}} \right)^{-\mu} \frac{\hat{\pi}_t^A}{v_t} + \frac{1 - \kappa + (\kappa - \bar{\kappa})(1 - G_t)}{\kappa} \hat{h}_t^A \right) - \frac{\eta G_t}{1 - \kappa} \left( \hat{h}_t^A \right)^{\kappa + 1} + \frac{1 - \bar{\kappa}}{1 - \kappa} g_t^N \hat{h}_t^A,
\]

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

\( g_t^N \) is still given by (48), \( \frac{\hat{\pi}_t^A}{v_t} \), \( H_t^P \) and \( \omega_t \) are now given by

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

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

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

which replace (45), (47) and (49).

One can then solve for a steady state of this system with \( G^* > 0 \) (and \( (g^N)^* > 0 \) so
that $n^* = 0$. (107) and (109) still apply, but (108) is replaced with

$$\hat{n}_A^* = \frac{\kappa \rho + (\theta - 1) \psi g^N}{\gamma 1 - \kappa + (1 - G^*) (\kappa - \bar{\kappa})},$$

(140)

and (111) with

$$f_{sp} (g^N) \equiv \frac{\rho + (\theta - 1) \psi g^N}{\psi H^P} \left( \frac{[(\hat{n}_A^*)]^{1-\kappa}}{\eta \kappa (G^*)^{\kappa-1} + \frac{1}{\gamma}} \right),$$

which is obtained by fixing $\hat{n}_t = 0$ in (139) using (107) and (140). 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 $\hat{n}_A^*$ 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 $H^P$ is identical to the equilibrium allocation once a subsidy to the use of intermediates $1/\sigma$ 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 inputs 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 $\Omega_t^A \equiv 1 - s^A_t$ and $\Omega_t^H$ similarly defined. In this situation, the law of motion for the private value of an automated firm, $V^A_t$, is still given by (19), for a non-automated firm it obeys:

$$r_t V^N_t = \pi^N_t - \Omega_t^A w_{Ht} h_t + \eta (G_t)^{\kappa} N_t^{\kappa} (h_t^A)^{\kappa} (V^A_t - V^N_t) + \dot{V}^N_t,$$

(141)

instead of (20), the first-order condition for automation is given by:

$$\kappa \eta (G_t)^{\kappa} N_t^{\kappa} (h_t^A)^{\kappa-1} (V^A_t - V^L_t) = \Omega_t^A w_{Ht},$$

(142)
instead of (21), while the free entry condition, when \( \gamma_t^N > 0 \), is given by
\[
\gamma_t N_t V_t^N = \Omega_t^H w_{Ht}, ~ (143)
\]
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 (135), (136), (137) and (138).

Using (137) and (143), we then get that \( \Omega_t^H \) must satisfy
\[
\Omega_t^H \dot{\zeta}_t = V_t^N, ~ (144)
\]
similarly, using (138) and (142), we get
\[
\Omega_t^A (\xi_t - \zeta_t) = V_t^A - V_t^L. ~ (145)
\]
Plugging (144) and (145) in (141), we get that
\[
r_t \zeta_t = \frac{\pi_t^N}{\Omega_t^H} - \frac{\Omega_t^A}{\Omega_t^A} w_{Ht} h_t + \eta (G_t) \tilde{K} N_t^\kappa (h_t^A)^\kappa \frac{\Omega_t^A}{\Omega_t^H} (\xi_t - \zeta_t) + \frac{\Omega_t^H}{\Omega_t^H} \dot{\zeta}_t + \dot{\zeta}_t. ~ (146)
\]
Similarly, using (145) and the difference between (19) and (141) gives:
\[
r_t (\xi_t - \zeta_t) = \frac{\pi_t^A - \pi_t^N}{\Omega_t^A} + w_{Ht} h_t - \eta (G_t) \tilde{K} N_t^\kappa (h_t^A)^\kappa (\xi_t - \zeta_t) + \frac{\Omega_t^A}{\Omega_t^A} (\xi_t - \zeta_t) + \dot{\zeta}_t - \dot{\zeta}_t. ~ (147)
\]
Combining (146) with (135), using (138) and (137) and the definition of \( \Omega_t^A \) and \( \Omega_t^H \), we get:
\[
\dot{s}_t^H = \frac{\gamma_t^N}{v_t} s_t^H - \left(1 - s_t^H\right) g_t^N + \frac{\tilde{\gamma}_t^A}{\kappa} \left((1 - s_t^A) (1 - \kappa) + (1 - s_t^H) \tilde{K} (1 - G_t) + \kappa G_t - 1)\right). ~ (148)
\]
Similarly combining (147) with the difference between (136) and (135) and using (137) gives:
\[
s_t^A \left( \frac{\tilde{h}_t^A}{v_t} \right)^{1 - \kappa} \frac{\pi_t^N}{\eta G_t^\kappa} = \frac{\gamma_t^N}{v_t} s_t^A - \frac{\tilde{\gamma}_t^A}{\kappa} s_t^A - \tilde{K} (1 - s_t^A) \tilde{h}_t^A \frac{1 - G_t}{G_t}. ~ (149)
\]
Therefore, in steady state, we have

\[ s^A_\infty = \frac{\tilde{\kappa} h^A_\infty (1 - G_\infty)}{\kappa \psi H^P + \tilde{\kappa} h^A_\infty (1 - G_\infty)} \geq 0. \]

Note from (149) that the share of automated products, \( s^A_t \), must always be non-negative, otherwise it cannot converge to a positive value, therefore \( s^A_t \geq 0 \) everywhere (and in fact \( > 0 \) if \( \tilde{\kappa} \neq 0 \)). Furthermore, if \( \tilde{\kappa} = 0 \), \( s^A_t = 0 \) 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).

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

\[ s^H_\infty = 1 - \frac{\gamma \tilde{h}^A_\infty (1 - \kappa) (1 - s^A_\infty)}{\kappa g^N_\infty + \gamma \tilde{h}^A_\infty (1 - \kappa) (1 - G_\infty - \kappa G_\infty)}. \]

In addition, knowing that \( s^A_t \geq 0 \), imposes that \( s^H_t > 0 \)—as \( s^H_t < 0 \) would lead to \( s^H_t < 0 \).

### 7.8.4 Transitional dynamics for the social planner case

Figure 23 plots the transitional dynamics for the optimal allocation in our baseline case (which features \( \tilde{\kappa} = 0 \)) and in the case where \( \tilde{\kappa} = 0 \) analyzed in Figure 7. 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 3.A and Figure 7.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

$$w_{Ht} \geq \gamma N_t V_t,$$

where

$$V_t = \min(\eta (N_t h_t^A)^\kappa, 1) V_t^A + (1 - \min(\eta (N_t h_t^A)^\kappa, 1)) V_t^N - w_{Ht} h_t^A.$$

is the expected value of a new firm. Since we used similar functional forms we have that $h_t^A$ obeys (21) unless $\kappa \eta^{\frac{1}{\kappa}} N_t (V_t^A - V_t^L) > w_{Ht}$, in which case $N_t h_t^A = \eta^{-\frac{1}{\kappa}}$. 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 (19). 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 \( g_t^N > 0 \).

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 \( \hat{h}_A^* = \min \left( \eta^{-\frac{1}{\kappa}}, \frac{\kappa}{1-\kappa} \right) \).

Therefore \( G^* < 1 \), if and only if \( \eta \left( \frac{\kappa}{1-\kappa} \right)^\kappa < 1 \). When \( G^* < 1 \), we will have that \( G_\infty = G^* < 1 \), so that, following Proposition 2,

\[
g_{wL}^\infty = \frac{1}{1 + \beta (\sigma - 1)} g_{wH}^\infty.
\]

On the contrary, if \( G^* = 1 \), then \( G_\infty = 1 \), and following Proposition 2, we get that

\[
g_{wL}^\infty = g_{wH}^\infty / \epsilon.
\]

Figure 24 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 \( \eta \left( \frac{\kappa}{1-\kappa} \right)^\kappa < 1 \), and the figure shows that the economy goes through three phases as in our baseline model.

7.10 Supply response in the skill distribution: details

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 (18) 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:

\[
L_t = t H \bar{j}_t.
\]
and the mass of high-skill labor is

\[ H_t = \overline{H} \left( 1 - \frac{1 + q}{q} \right) \leq \overline{H}. \]  \hspace{1cm} (151)

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

\[ \overline{j}_t = \left( \frac{q}{1 + q} \frac{l_w_{Lt}}{w_{Ht}} \right)^q. \]  \hspace{1cm} (152)

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

**7.10.1 Asymptotic growth rates**

We consider processes \( (N_t, G_t, H_t^P) \) such that \( g_t^N, G_t \) and \( H_t^P \) admit strictly positive limits. Plugging (152) and (150) in (10), we get:

\[ \frac{w_{Ht}}{w_{Lt}} = l \left( \frac{1 - \beta}{\beta} \frac{\overline{H}}{H_t^P} \right)^q \frac{G_t + (1 - G_t) (1 + \varphi w_{Lt}^{-1})^{-\mu}}{G_t (1 + \varphi w_{Lt}^{-1})^{-1} + (1 - G_t) (1 + \varphi w_{Lt}^{-1})^{-\mu}} \right)^{\frac{1}{1 + q}}, \]  \hspace{1cm} (153)
which together with (11) determines \( w_{HT} \) and \( w_{Lt} \) for given \((N_t, G_t, H^P_t)\). From then on the reasoning follows that of Appendix 7.2.1. First, we derive that \( w_{L\infty} > 0 \), such that \( g^{w_L}_{\infty} = g^{GDP}_{\infty} = \psi g^N_{\infty} \), and that we must have \( g^{w_L}_{\infty} < g^{w_H}_{\infty} \), such that \( j_{\infty} = 0 \). Second, we study the asymptotic behavior of \( w_{Lt} \) both when \( \epsilon < \infty \) and when \( \epsilon = \infty \).

**Case with \( \epsilon < \infty \).** Plugging (153) in (11) gives \( w_{Lt} \) in function of \( N_t, G_t \) and \( H^P_t \):

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

which replaces (59). It is direct that when \( G_{\infty} < 1 \), we obtain (50). In this case, we further have

\[
g_\infty^{-1} = q (g^{w_L}_{\infty} - g^{w_H}_{\infty}) = -\frac{q \beta (\sigma - 1)}{1 + q + \beta (\sigma - 1)} g^{GDP}_{\infty} . \tag{155}
\]

**Case with \( \epsilon = \infty \).** In this case, (154) becomes

\[
w_{Lt} = \frac{\sigma - 1}{\sigma} \beta \left( \frac{1 + q}{q} \right)^{(1 - \beta)q} \frac{H^P_t}{H^T}^{1 - \beta} \left( \frac{H^P_t}{H^T} \right)^{1 - \frac{1}{1+q}} \left( G_t (\varphi w_{Lt}^{-1} (n - 1) + (1 - G_t)) \right)^{1 - \beta} , \text{ if } w_{Lt} > \varphi^{-1} ,
\]

\[
w_{Lt} = \frac{\sigma - 1}{\sigma} \beta \left( \frac{1 + q}{q} \right)^{(1 - \beta)q} \frac{H^P_t}{H^T}^{1 - \beta} \left( \frac{H^P_t}{H^T} \right)^{1 - \frac{1}{1+q}} N^{1 - \frac{1}{1+q}} , \text{ if } w_{Lt} < \varphi^{-1} .
\]

Once again, following the steps of Appendix 7.2.1, we get that if \( G_{\infty} < 1 \), (50) applies (and accordingly we also get (155)).

### 7.10.2 Dynamic system

It is convenient to redefine \( n_t \equiv N_t^{1 + \beta (\sigma - 1)} \), we can then write the entire dynamic system as a system of differential equations in \((n_t, G_t, \hat h^A_t, \chi_t)\) with two auxiliary variables \( \omega_t \) and \( \hat j_t \equiv j_t n_t^{1 + \beta (\sigma - 1)} \). Equations (31) is now given by

\[
\dot n_t = -\frac{\beta}{1 - \beta} \frac{1 + q}{1 + q + \beta (\sigma - 1)} g^N_t n_t ,
\]
(32), (39), (40), (45), (47) still apply and equation (48) as well provided that \( H \) is replaced by \( H_t \) given by (151). \( \omega_t \) is implicitly defined by:

\[
\omega_t = \left( \left( \frac{\sigma-1}{\sigma} \right)^{\frac{1+q}{1+q}} \beta^{1+\frac{q}{1+q}} \left( 1 - \beta \right)^{\frac{1+q}{q}} \frac{H^p}{1+q+H} \left( G_t \left( 1 + \varphi \left( \omega_t n_t \right)^{-\frac{1}{q}} \right)^{\mu-1} + 1 - G_t \right) \right)^{\frac{\beta(1-\sigma)}{1+q+\beta(\sigma-1)}} \times \left( G_t \left( \varphi + \left( \omega_t n_t \right)^{\frac{1}{q}} \right)^{\mu} + 1 - G_t \right) \omega_t \right)^{(1+q)-1},
\]

which replaces (49) and is a rewriting of (154) and \( \hat{j}_t \) is given by:

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

which is derived using (152) and (153).

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

### 7.11 Machines as a capital stock

#### 7.11.1 Set-up

To avoid repetitions, we already include the taxes of section 4.2, namely, we assume that there is a tax \( \tau_m \) on the rental rate of equipment and a tax \( \tau_h \) 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:

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

The Euler equation (24) still applies and the capital accumulation equation is given by (30). The unit cost of intermediate input \( i \) is now given by

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

(157)
instead of (4) where \( \varphi \equiv \tilde{\varphi}^\epsilon \left( \frac{\beta_2}{\beta_4} \left( 1 - \beta_4 \right) \right)^{1-\epsilon} \). Define \( \mu \equiv \beta_1 (\sigma - 1) / (\epsilon - 1) \), we can then derive the isocost curve as:

\[
N \frac{1}{\sigma - 1} \frac{\sigma}{\sigma - 1} \tilde{\varphi}^\epsilon \left( \frac{\beta_2}{\beta_4} \left( 1 - \beta_4 \right) \right)^{1-\epsilon} \mu = 1.
\]

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

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

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:

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

### 7.11.2 Effect of technology on wages

First note that one can rewrite (159)

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

115
where we defined

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

In (161), 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 (161) in (158) 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 10.** Consider the equilibrium $(w_L, w_H)$ determined by equations (161) and (158). 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 (158) as:

$$N^{\frac{1}{\sigma}} \frac{\sigma}{\sigma - 1} \beta_3 \bar{r} \beta_3 \left( \frac{w_H}{w_L} \right)^{\beta_3} \left( \frac{w_H}{w_L} \right)^{\beta_3} \left( \varphi \left( \left( (1 + \tau_m) \bar{r} \right)^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right)^{\mu} \left( 1 - G \right) w_L^{\beta_3 (1-\sigma)} = 1.$$ (162)

Using that (161) establishes $\frac{w_H}{w_L}$ as an increasing function of $w_L$ otherwise independent of $N$, we get that (162) implies that $w_L$ and therefore $w_H/w_L$ (when $G > 0$) and $w_H$ itself must increase in $N$.

(161) 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 (158) the direct effect of an increase in $G$ is to decrease the LHS (because $\left( \varphi \left( \left( (1 + \tau_m) \bar{r} \right)^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right)^{\mu} > w_L^{\beta_3 (1-\sigma)}$), in addition an increase in $G$ would reduce $w_L$ which further reduces the LHS. To maintain
the inequality, it must be that \( w_H \) increases. Therefore in this case too, \( w_H \) increases in \( G \) and so does \( w_H / w_L \).

This model is isomorphic to the previous one when \( \beta_4 = \beta_3 = 0 \) (with \( \varphi ((1 + \tau_m) \tilde{r})^{1-\epsilon} \) replacing \( \varphi \)), therefore the impact of a change of \( G \) on \( w_L \) is also ambiguous.

The labor share is now given by

\[
LS = \frac{w_L L + w_H H}{Y + (1 + \tau_a) w_H (H - H^P)}.
\]

As before profits are a share \( \frac{1}{\sigma} \) of output so that

\[
Y = \frac{\sigma}{\sigma - 1} \left( w_L L + w_H H^P + \tilde{r} K + T_m \right), \tag{163}
\]

where \( T_m \) denotes the tax proceeds from the tax on equipment. We have

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

Then, we obtain:

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

Assume that \( H = H^P \), then we get that

\[
LS = \frac{\sigma - 1}{\sigma} \left( 1 + \frac{\tilde{r} K + T_m}{w_L L + w_H H} \right)^{-1}.
\]
Using (159), (160), (164) and the definition of $\Phi$, we obtain:

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

This expression is increasing in $\Phi$. From (161), $\Phi$ 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 (161) as:

$$\frac{w_H H^P}{w_L L} = \frac{\beta_3}{\beta_2 + \beta_1} \frac{1}{\beta_2 + \beta_1} \frac{G \Phi (\Phi + 1)^{\mu - 1}}{G (\Phi + 1)^{\mu - 1} + (1 - G)}.$$

We have already derived that an increase in $G$ increases $w_H/w_L$, therefore, this expression shows that it will increase $\frac{G \Phi (\Phi + 1)^{\mu - 1}}{G (\Phi + 1)^{\mu - 1} + (1 - G)}$. We can then rearrange terms in (166) and write:

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

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

### 7.11.3 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 (156)). 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 11.** Consider four processes \([N_t]_{t=0}^\infty, [G_t]_{t=0}^\infty, [H_t^P]_{t=0}^\infty, \tilde{r}_t \]
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 \( w_{Ht} \) and output \( Y_t \) are

\[
g_\infty^{w_H} = g_\infty^Y = g_\infty^N/[(\sigma - 1)(\beta_2 + \beta_1 \beta_4)],
\]

and the asymptotic growth rate of low-skill wages is

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

**Proof.** For simplicity we assume that the limits \( g_\infty^{w_H}, g_\infty^{w_L} \) and \( g_\infty^Y \) exist (although we could show that formally as we did in Appendix 7.2.1). Suppose that \( g_\infty^{w_L} \leq \beta_4 g_\infty^{w_H} \). Then \( \Phi_t \) must either tend toward a positive constant or toward 0, in either case (161) implies that \( g_\infty^{w_L} = g_\infty^{w_H} \), which is a contradiction as \( \beta_4 < 1 \). Hence it must be that \( g_\infty^{w_L} > \beta_4 g_\infty^{w_H} \), which ensures that \( \Phi_t \to \infty \). Using this in (158), we obtain:

\[
w_{Ht}^{\beta_2 + \beta_1 \beta_4} \to \frac{\sigma - 1}{\sigma} \frac{\beta_1 \beta_4^{\beta_2 \beta_3}}{(1 + \tau_m)(1 - \beta_4\beta_1 - \beta_3)} \frac{G_\infty \varphi}{\tilde{r}_t^{\beta_2 + (1 - \beta_4)\beta_1}} N_t^{-1/(\sigma - 1)}.
\]

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

Moreover (161) now implies

\[
\frac{w_{Ht} H_t^P}{w_{Lt} L} \to \frac{G_\infty (\beta_2 + \beta_1 \beta_4) \Phi_t^\mu}{\beta_1 (1 - G_\infty)},
\]

\[
\implies w_{Lt}^{1 + \beta_1 (\sigma - 1)} \to \frac{\beta_1 (1 - G_\infty) (1 + \tau_m)(1 - \beta_4 \beta_1 - \beta_3)}{G_\infty (\beta_2 + \beta_1 \beta_4) \varphi^\mu L} \frac{H_t^P}{w_{Ht}^{1 + \beta_4 \beta_1 (\sigma - 1)}}
\]

which implies (168). Since \( \frac{1 + (\sigma - 1)\beta_1 \beta_4}{1 + (\sigma - 1)\beta_1} > \beta_4 \), we verify that \( g_\infty^{w_L} > \beta_4 g_\infty^{w_H} \).

**7.11.4 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: \( \tilde{V}_t^A, \tilde{V}_t^N, \tilde{N}_t^A, \tilde{N}_t^N, \tilde{h}_t^A, \tilde{c}_t \) and \( \tilde{\nu}_t \). We also introduce \( \hat{Y}_t \equiv Y_t N^{-\psi} \) and \( \hat{K}_t \equiv K_t N^{-\psi} \). Finally we now define

\[
n_t \equiv N_t \left[ 1 - \frac{\beta_4}{1 + \beta_1 (\sigma - 1) \beta_2 + \beta_1 \beta_4} \right]^{1/\beta_1} \]

and

\[
\omega_t \equiv \left( \frac{w_L}{\tilde{w}_H^{1-\beta_4}} \right)^{\beta_1 (1-\sigma)} \left( \frac{1-\beta_4}{w_H^{1-\beta_4}} \right)^{\beta_1 (\sigma - 1)} = \omega_t n_t.
\]

so that

\[
(\omega_t n_t)^{1/\beta_1} = \omega_t n_t.
\]

The transitional dynamics can then be expressed as a system of differential equations in \( x_t \equiv (n_t, G_t, \hat{K}_t, \hat{h}_t^A, \hat{c}_t, \tilde{\nu}_t) \) where the first three variables are state variables and the last three control variables.

Equation (41) still applies, therefore, we get using (157) that

\[
\pi_t^A = \frac{(\sigma - 1)^{\sigma - 1}}{\sigma^\rho} \left( \phi \left( \frac{1}{(1 + \tau_{m}) \tilde{r}_t} \right)^{1-\beta_4} \frac{w_H^{\beta_4}}{w_H^{1-\beta_4}} \right)^{1-\epsilon} + w_L^{1-\epsilon} \left( \frac{w_H^{\beta_4}}{w_H^{1-\beta_4}} \right)^{1-\rho} Y_t.
\]

We can rewrite this as

\[
\pi_t^N = \frac{(\sigma - 1)^{\sigma - 1}}{\sigma^\rho} \left( \phi \left( \frac{1}{(1 + \tau_{m}) \tilde{r}_t} \right)^{1-\beta_4} \frac{w_H^{\beta_4}}{w_H^{1-\beta_4}} \right)^{1-\sigma} \left( \phi (1 + \tau_{m})^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{1/\beta_1} \right)^{1-\rho} \hat{Y}_t.
\]

Equation (170) is now replaced by

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

(32) still applies and so does (33). Because of the automation tax (34) is replaced by

\[
(r_t - (\psi - 1) g_t^N) \hat{V}_t^N = \hat{N}_t^N + \eta \frac{\hat{h}_t^A (\hat{h}_t^A)^{\kappa}}{\hat{V}_t^A - \hat{V}_t^N} - (1 + \tau_{a}) \hat{\nu}_t \hat{h}_t + \hat{V}_t^N \]

(172)
and (35) by
\[ \kappa \eta G_t^\kappa \left( \dot{h}_t^A \right)^{\kappa-1} \left( \dot{V}_t^A - \dot{V}_t^N \right) = (1 + \tau_a) \dot{v}_t. \] (173)
Combining (170), (172), (173) and (23) with equality, we now obtain:
\[ \dot{v}_t = \dot{v}_t \left( \tilde{r}_t - \Delta - (\psi - 1) g_t^N - \gamma \omega t (\varphi (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega t n_t)^{1/\mu}) \right) - \mu \frac{\pi_t^A}{\dot{v}_t} - \gamma (1 + \tau_a) \frac{1 - \kappa}{\kappa} \dot{h}_t^A. \] (174)
Following the same steps as those used to derive (39), we now obtain:
\[ \dot{h}_t^A = \gamma \frac{\dot{h}_t^A}{1 - \kappa} \left( \omega t n_t \left( \varphi (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega t n_t)^{1/\mu} \right) \right) - \mu \frac{\pi_t^A}{\dot{v}_t} + (1 + \tau_a) \frac{1 - \kappa}{\kappa} \dot{h}_t^A \]
\[ = \frac{\kappa \eta G_t^\kappa}{\left( 1 - \kappa \right) (1 + \tau_a)} \left( 1 - \omega t n_t \left( \varphi (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega t n_t)^{1/\mu} \right) \right) - \mu \frac{\pi_t^A}{\dot{v}_t} + \eta G_t^\kappa \left( \dot{h}_t^A \right)^{\kappa+1} \]
\[ + \frac{\kappa}{1 - \kappa} \left( \eta \left( \dot{h}_t^A \right)^{\kappa+1} G_t^{\kappa-1} (1 - G_t) - g_t^N \dot{h}_t^A \right). \] (175)
Further, (24) still applies and we can rewrite it as:
\[ \dot{c}_t = \frac{\dot{c}_t}{\vartheta} \left( \tilde{r}_t - (\rho + \Delta + \theta g_t^N) \right). \] (176)
Finally, we can rewrite (30) as
\[ \dot{K}_t = \dot{Y}_t - \dot{c}_t - (\Delta + \psi g_t^N) \dot{K}_t \] (177)
Equations (171), (32), (174), (175), (176) and (177) form a system of differential equations which depend on \( \dot{Y}_t, \pi_t^A, \tilde{r}_t \) and \( g_t^N \).
(158) implies
\[ \frac{\sigma}{\sigma - 1} \frac{\nu^{\beta_2 + \beta_4 \beta_1 \beta_3 + \beta_1 (1 - \beta_4)}}{\beta_1 \beta_2 \beta_3} \left( G \left( \varphi (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega t n_t)^{1/\mu} \right) \right)^{1 - \sigma} = 1, \]
so that
\[
\tilde{r} = \left[ \frac{\sigma - 1}{\sigma} \beta_1 \beta_2 \beta_3 \right]^{1-\beta_4} G \left( \varphi (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\beta}} \right)^\mu + (1 - G) \omega_t n_t \frac{1}{\beta_3 + \beta_2 (1-\sigma)}
\]

which defines \( \tilde{r} \) as a function of \( x_t \) and \( \omega_t \). (160) can be written as:
\[
H_t^P = \tilde{r}_t \tilde{K}_t \frac{G_t \left( \beta_2 \left[ \frac{\beta_1 \beta_4 (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)}}{(\omega_t n_t)^{\frac{1}{\beta}} + (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)}} \right] \left( \varphi (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\beta}} \right)^\mu + \beta_2 (1 - G_t) \omega_t n_t \right)}{\tilde{v}_t}
\]

which gives, together with (178), \( H_t^P \) as a function of \( x_t \) and \( \omega_t \). \( g_t^N \) still obeys (48), which then defines it as a function of \( x_t \) and \( \omega_t \).

Combine (163), (159), (160) and (164) to obtain:
\[
\frac{Y}{w_H^PH^P} = \frac{\sigma}{\sigma - 1} \left[ G_t \left( \varphi ((1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\beta}} \right)^\mu + (1 - G) \omega_t n_t \right] \tilde{v}_t H_t^P
\]

which we can rewrite as
\[
\hat{Y}_t = \frac{\sigma}{\sigma - 1} \left[ G_t \left( \varphi ((1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\beta}} \right)^\mu + (1 - G) \omega_t n_t \right] \tilde{v}_t H_t^P
\]

This expression, with the previous equations, gives \( \hat{Y}_t \) as a function of \( x_t \) and \( \omega_t \). (169)

then ensures that \( \hat{\pi}_t^A \) is defined as a function \( x_t \) and \( \omega_t \).

Finally, from (159) we obtain:
\[
\omega_t = \left[ \frac{G_t \left( \varphi ((1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\beta}} \right)^\mu + (1 - G) \omega_t n_t \right] \tilde{v}_t H_t^P}{G_t \left( \beta_2 + \frac{\beta_1 \beta_4 (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)}}{(\omega_t n_t)^{\frac{1}{\beta}} + (1 + \tau_m)^{(1-\beta_4)(1-\epsilon)}} \right) \left( \varphi ((1 + \tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\beta}} \right)^\mu + \beta_2 (1 - G) \omega_t n_t}
\]

122
which implicitly defines $\omega_t$ as a function of $x_t$. Hence, together with (178), (179), (48), (180), (169) and (181), the system formed by (171), (32), (174), (175), (176) and (177) describes the dynamic equilibrium. We then obtain

**Proposition 12.** Assume that

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

(182)

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

**Proof.** As before, we directly get that in a steady-state with $g^{N*} > 0$, we must have $n^* = 0$. (181) then implies that $\omega^*$ is a constant defined by

$$
\omega^* = \left[ \frac{\left( \frac{\hat{v}^*}{\tilde{r}^*} \right)^{1 - \beta_1} \frac{H^P^* \beta_1 (1 - G^*) (1 + \tau_m)^{(1 - \beta_4)(1 - \sigma)} G^* (\beta_2 + \beta_1 \beta_4) \varphi^\mu}{\bar{L}}}{\bar{\gamma}^*} \right]^{\beta_1 (1 - \sigma) \frac{1}{1 + \beta_1 (\sigma - 1)}}.
$$

This guarantees that in such a steady-state, $w_{LT} \sim \omega^* \frac{1}{\bar{r}^* (\sigma - 1)} \frac{1}{\beta_4 \beta_3} \frac{1}{\bar{\gamma}^* (\beta_2 + \beta_1 \beta_4) N_t}$ such that $g_{LT}^{w} = \frac{1}{\bar{r}^* (\sigma - 1)} \frac{1}{\beta_4 \beta_3} \frac{1}{\bar{\gamma}^* (\beta_2 + \beta_1 \beta_4) N_t}$ as stipulated in Proposition 11.

In addition, (176) implies that in steady-state,

$$
\tilde{r}^* = \rho + \Delta + \theta \psi g^{N*}.
$$

(183)

(179) implies that

$$
H^{P*} = \frac{\tilde{r}^* \hat{K}^*}{\tilde{v}^*} \frac{\beta_2 + \beta_1 \beta_4}{\beta_3 + \beta_1 (1 - \beta_4)}.
$$

(184)

Then (180) implies that

$$
\tilde{Y}^* = \frac{\sigma}{(\sigma - 1) (\beta_2 + \beta_1 \beta_4)} \tilde{v}^* H^{P*}.
$$

(185)

We then get that (169) implies that

$$
\frac{\hat{\pi}^{A*}}{\hat{v}^*} = \frac{1}{\sigma^{1 - \sigma - 1}} \left( \beta_1 \beta_2 \beta_3 \right)^{\sigma - 1} \left( \tilde{r}^* (1 - \beta_4) \beta_1 \beta_2 + \beta_1 \beta_4 \beta_3 (1 + \tau_m)^{(1 - \beta_4) \beta_1} (1 + \tau_m) \left( \beta_2 + \beta_1 \beta_4 \right) \right)^{1 - \sigma} \varphi^\mu H^{P*}.
$$

(186)
(178) gives
\[
\hat{r}^* = \left[ \frac{\sigma - 1}{\sigma} \beta_1 \beta_2 \beta_3 \beta_4 \left( G^* \varphi^\mu \right) \frac{1}{1 + \tau_a} \right]^{\frac{1}{\beta_3 + \beta_4 (1 - \beta_1)}}.
\]

Therefore (186) simplifies into
\[
\frac{\hat{r}^*}{\hat{v}^*} = \frac{\psi H^P*}{G^*},
\]
just as in the baseline model. Then (174) and (183) together imply that
\[
\hat{h}^A* = \frac{\kappa}{\gamma (1 + \tau_a) (1 - \kappa)} \left( \rho + ((\theta - 1) \psi + 1) g^{N*} \right).
\]

This defines \( \hat{h}^A* \) as an increasing function of \( g^{N*} \). Further, in steady-state \( G^* \) still obeys (107) and \( H^P* \) obeys (109), which imply that \( G^* \) and \( H^P* \) also be defined as function of \( g^{N*} \).

(188), (175), (107), (189) then lead to
\[
\frac{1 - \kappa \gamma G^* (1 + \tau_a)}{\gamma H^P*} \left( \frac{1}{\kappa \eta G^*_t} \left( \hat{h}^A* \right)^{1 - \kappa} + \frac{1}{\gamma} \right) = 1,
\]
which up to the term \( 1 + \tau_a \) is the same as (111) 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 (182) is satisfied. As (190), (107), (109) and (189) are independent of \( \tau_m \), so are \( g^{N*} \), \( \hat{h}^A* \) (now given by (189)), \( G^* \) (given by (107)) and \( H^P* \) (given by (109)).

We further obtain \( \hat{r}^* \) through (183), which must be independent of \( \tau_m \) as well. We then get \( \hat{v}^* \) through (187) as
\[
\hat{v}^* = \left[ \frac{\sigma - 1}{\sigma} \beta_1 \beta_2 \beta_3 \beta_4 \left( G^* \varphi^\mu \right) \frac{1}{1 + \tau_m} \right]^{\frac{1}{\beta_3 + \beta_4 (1 - \beta_1)}}.
\]

We then get \( \hat{K}^* \) through (184) and \( \hat{c}^* \) from (177) which, using (185), implies:
\[
\hat{c}^* = \frac{\sigma}{(\sigma - 1) (\beta_2 + \beta_1 \beta_4)} \hat{v}^* H^P* - (\Delta + \psi g^{N*}) \hat{K}^*.
\]

Further if \( \tau_a = \tau_m = 0 \), \( g^{N*}, G^*, \hat{h}^A* \) 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 4 extends to this case.

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 \( \hat{Y}_i = \{\hat{Y}_{i,t}\}_{t=1}^{T_i} \) for each of these four paths from 1963 and until 2007 for the labor share, skill-premium, and GDP per hour, and 2000 for equipment over GDP (due to data limitations from Cummins and Violante, 2002). We let \( \hat{Y}(b) = \{\hat{Y}_i(b)\}_{i=1}^4 \) as the combined vector of these paths and make explicit the dependency on the parameters \( b \). \( Y \) is the corresponding vector of actual values. We then solve:

\[
\min_b (\log(\hat{Y}(b)) - \log(Y))^\prime W (\log(\hat{Y}(b)) - \log(Y)),
\]

where \( W \) is a diagonal matrix of weights. In a previous version of the paper (Hémous and Olsen, 2016) we articulated a stochastic version of our model by introducing auto-correlated 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 (2012). However, since our estimation requires a skill-ratio both before and after the period 1963-2007 we match the observed path of the log of the skill-ratio to a “generalized”
logistical function of the form:

\[
\frac{\alpha}{1 + exp \left( \frac{\mu - t}{s} \right)} + \beta,
\]

where \((\alpha, \beta, \mu, s)\) 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 directly from Acemoglu and Autor (2012) specifically the data underlying their Figures 1 and 2 from David Autor’s website. The labor share is taken from Koh, Santaeulàlia-Llopis and Zheng (2016). 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 Krusell et al. (2000) and use quality-adjusted price indices of equipment from Cummins and Violante (2002) who update the series from Gordon (1990). We combine two different series. First, we use NIPA data on private investment in equipment excluding transportation equipment (Tables 1.5 and 5.3.5. from NIPA). We iteratively construct an index for the stock of private real capital equipment by assuming a depreciation rate of 12.5 per cent (as Krusell et al., 2000) and using the price index for private equipment from Cummins and Violante (2002). We start this approach in 1947 but only use the stock from 1963 onwards. We combine this with the growth rate of real private GDP to get an index for equipment over GDP. We match this index to the NIPA private equipment capital stock (excluding transportation) over (private) GDP number for 1963 to get a series in absolute value. To this, we add software, but following the suggestion of Cummins and Violante (2002), we use the NIPA data on the stock of software over GDP (table 2.1 from NIPA). We add these two values to get our combined stock of equipment (+software) over GDP.