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

David Hémous and Morten Olsen

July 2020

B. Online Appendix

B1. Additional results for Section I

B1.1 Proof of Proposition 1 in the perfect substitute case: $\epsilon = \infty$

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

$$w_L = \frac{\sigma-1}{\sigma}\beta\left(\frac{H^P}{L}\right)^{1-\beta} N^{\frac{1}{\sigma-1}}, \tag{B1}$$

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

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

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

Case iii) $w_L > \widetilde{\varphi}^{-1}$ and all automated firms use machines only, in that case, we get that (A2) is replaced by

$$w_L = \frac{\sigma-1}{\sigma}\beta\left(\frac{H^P}{L}\right)^{(1-\beta)} N^{\frac{1}{\sigma-1}}(1-G)^{1-\beta}\left(G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)} + 1 - G\right)^{\frac{1}{\sigma-1}-(1-\beta)}, \tag{B2}$$

and the skill premium obeys:

$$\frac{w_H}{w_L} = \frac{1-\beta}{\beta}\frac{L}{H^P}\frac{G(w_L\widetilde{\varphi})^{\beta(\sigma-1)} + 1 - G}{1-G}. \tag{B3}$$

One can rewrite (B2) as

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

The left-hand side increases in w_L and the right-hand side decreases in w_L, so this expression defines w_L uniquely. Moreover w_L is greater than $\widetilde{\varphi}^{-1}$ if and only if

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

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

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

Comparative statics of w_L with respect to G. Furthermore, (B2) 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 (B2), one gets:

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

where

$$Den \equiv 1 - \beta\frac{G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)}}{G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)} + 1 - G} + \frac{(1-\beta)\beta(\sigma-1)G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)}}{G(\widetilde{\varphi}w_L)^{\beta(\sigma-1)} + 1 - G}. \tag{B4}$$

We have

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

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

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

Log-differentiating (B2) with respect to both N and G, one gets:

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

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

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

Therefore w_L increases.

Log-differentiating (B3), we get:

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

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

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

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

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

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

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

B1.2 Comparative statics with respect to $\widetilde{\varphi}$

We now look at the effect of an increase in machine's productivity $\widetilde{\varphi}$ (which up to some relabeling is equivalent to a decline in the price of machines). We focus on the case $\epsilon < \infty$, so that an increase in $\widetilde{\varphi}$ is equivalent to an increase in φ. To look at its effect on low-skill wages, log-differentiate (A2)

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

where Den is still given by (B4). We then get that $\frac{\partial w_L}{\partial \varphi} < 0$ if $\psi \leq 1$ (that is $(1-\beta)(\sigma-1) > 1$), in which case $\frac{\partial w_L}{\partial G} < 0$. Provided that $\psi > 1$, we have that

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

Using (A3), we get

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

Note that

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

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

B2. Proofs of the asymptotic results

B2.1 Asymptotic results when $G_\infty \notin (0,1)$

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

PROPOSITION B.1: *Consider three processes $[N_t]_{t=0}^{\infty}$, $[G_t]_{t=0}^{\infty}$ and $[H_t^P]_{t=0}^{\infty}$ where $(N_t, G_t, H_t^P) \in (0,\infty)\times[0,1]\times(0,H]$ for all t. Assume that G_t, g_t^N and H_t^P all admit limits G_∞, g_∞^N and H_∞^P with $g_\infty^N > 0$ and $H_\infty^P > 0$.*

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

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

$$g_\infty^{w_L} = g_\infty^{Y}/\epsilon. \tag{B5}$$

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

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

$$g_\infty^{w_L} = g_\infty^{w_H} = g_\infty^{Y} = g_\infty^{N}/(\sigma-1). \tag{B6}$$

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

Subcase with $\epsilon < \infty$. With $G_\infty = 1$, equation (A2) still implies that w_{Lt} is unbounded and gives:

(B7)
$$w_{Lt} \sim \left(\left(\frac{\sigma-1}{\sigma}\beta\right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L} \varphi^{\mu(\psi-1)}\right)^{\frac{1}{\epsilon}} N_t^{\frac{\psi}{\epsilon}} \left(\varphi^{\mu-1} + (1-G_t)\, w_{Lt}^{(\epsilon-1)(1-\mu)}\right)^{\frac{1}{\epsilon}}.$$

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

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

Yet, this implies

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

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

Consider now the possibility that $\lim\,(1-G_t)\, w_{Lt}^{(\epsilon-1)(1-\mu)} = 0$, then (B7) implies

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

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

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

$$w_{Lt_m} \sim \left(\left(\frac{\sigma - 1}{\sigma} \beta \right)^{\frac{1}{1-\beta}} \frac{H_\infty^P}{L} \varphi^{\mu(\psi-1)} \left(\varphi^{\mu-1} + \lambda_1 \right) \right)^{\frac{1}{\epsilon}} N_{t_m}^{\frac{\psi}{\epsilon}}. \tag{B8}$$

This leads to

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

which is only possible if $\lim \left(1 - G_t\right) N_t^{\frac{\psi}{\epsilon}(\epsilon-1)(1-\mu)} > 0$. We denote such a limit by λ. Then (B7) leads to

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

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

Subcase with $\epsilon = \infty$. Low skill wages are defined as described in Online Appendix B1.1. With $G_\infty = 1$ and knowing that $\lim \left(1 - G_t\right) N_t^\psi$ exists and is finite, (B1) implies that w_{Lt} must be bounded weakly above $\widetilde{\varphi}$ in the long-run. As a result, (B2) leads to

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

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

Case where $G_\infty = 0$ (Part B). If $\lim G_t = 0$ then (A2) implies that for $\epsilon < \infty$:

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

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

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

$$w_{Lt} \sim \frac{\sigma-1}{\sigma}\beta\left(\frac{H_\infty^P}{L}\right)^{(1-\beta)} N_t^{\frac{1}{\sigma-1}}\left(G_t w_{Lt}^{\beta(\sigma-1)}\varphi^\mu + 1\right)^{\frac{1}{\sigma-1}-(1-\beta)}. \tag{B9}$$

Note that if $\epsilon = \infty$, then we must be in case iii) when $G_\infty = 0$ and (B2) also directly implies (B9) (as $\varphi^\mu = \widetilde{\varphi}^{\beta(\sigma-1)}$ in that case).

Assume that $\lim_{t\to\infty} G_t N_t^\beta = \lambda$ exists and is finite. Then (B9) implies:

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

which implies that $\lim_{t\to\infty} w_{Lt}N_t^{-\frac{1}{\sigma-1}}$ exists and is finite as well. Therefore one gets that $g_\infty^{w_L} = g_\infty^N/(\sigma-1)$. Using (A1) then immediately implies (B6). ■

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

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

B2.2 Sufficient conditions for $G_\infty \in (0,1)$

We prove the following Lemma:

LEMMA B.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_tN_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 $0 \le \nu_{2,t} < 1$ be the fraction of new products introduced at time t that are initially automated. Then

$(\dot{G_t N_t}) = \nu_{1,t}(1-G_t)N_t + \nu_{2,t}\dot{N}_t$ such that $\dot{G}_t = \nu_{1,t}(1-G_t) - (G_t - \nu_{2,t})g_t^N$. First assume that $G_\infty = 1$, then if $\nu_{1,t} < \bar{\nu}_1 < \infty$ and $\nu_{2,t} < \bar{\nu}_2 < 1$, we get that $\dot{G}_t$ must be negative for sufficiently large t, which contradicts the assumption that $G_\infty = 1$. Similarly if $G_\infty = 0$, then having $\nu_{1,t} > \underline{\nu}$ for all t, gives that $\dot{G}_t$ must be positive for sufficiently large t, which also implies a contradiction. Hence a limit must have $0 < G_\infty < 1$.

B3. An endogenous supply response in the skill distribution: static model

We present here an extension of the baseline model with an endogenous supply response in the skill distribution. Specifically, let there be a unit mass of heterogeneous individuals, indexed by $j \in [0,1]$ each endowed with $l\bar{H}$ units of low-skill labor and $\Gamma(j) = \bar{H}\frac{(1+q)}{q}j^{1/q}$ units of high-skill labor (the important assumption here is the existence of a fat tail of individuals with low ability). The parameter $q > 0$ governs the shape of the ability distribution with $q \to \infty$ implying equal distribution of skills and $q < \infty$ implying a ranking of increasing endowments of high-skill on $[0, \bar{H}(1+q)/q]$.

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

$$L_t = l\overline{H}\bar{j}_t, \tag{B10}$$

and the mass of high-skill labor is

$$H_t = \overline{H}\left(1 - \bar{j}_t^{\frac{1+q}{q}}\right) \leq \overline{H}. \tag{B11}$$

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

$$\bar{j}_t = \left(\frac{q}{1+q}\frac{lw_{Lt}}{w_{Ht}}\right)^q. \tag{B12}$$

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

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

PROPOSITION B.2: *Consider three processes $[N_t]_{t=0}^\infty$, $[G_t]_{t=0}^\infty$ and $[H_t^P]_{t=0}^\infty$ where*

$(N_t, G_t, H_t^P) \in (0,\infty) \times (0,1) \times (0,H]$ for all t. Assume that G_t, g_t^N and H_t^P all admit limits G_∞, g_∞^N and H_∞^P with $G_\infty \in (0,1)$, $g_\infty^N > 0$ and $H_\infty^P > 0$. Then the asymptotic growth of high-skill wages w_{Ht}, output Y_t and low-skill wages are:

$$g_\infty^{w_H} = g_\infty^Y = g_\infty^N \text{ and } g_\infty^{w_L} = \frac{1+q}{1+q+\beta(\sigma-1)} g_\infty^Y. \tag{B13}$$

Proof. We consider processes $\left(N_t, G_t, H_t^P\right)$ such that g_t^N, G_t and H_t^P admit strictly positive limits. Plugging (B12) and (B10) in (7), we get:

$$\frac{w_{Ht}}{w_{Lt}} = l \left(\frac{1-\beta}{\beta} \frac{\overline{H}}{H_t^P} \left(\frac{q}{1+q} \right)^q \frac{G_t + (1-G_t)\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-\mu}}{G_t\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-1} + (1-G_t)\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{-\mu}} \right)^{\frac{1}{1+q}}, \tag{B14}$$

which together with (6) determines w_{Ht} and w_{Lt} for given $\left(N_t, G_t, H_t^P\right)$. From then on the reasoning follows that of Appendix A2. First, we derive that $w_{L\infty} > 0$, such that $g_\infty^{w_H} = g_\infty^{GDP} = \psi g_\infty^N$, and that we must have $g_\infty^{w_L} < g_\infty^{w_H}$, such that $\bar{j}_\infty = 0$. Second, we study the asymptotic behavior of w_{Lt} both when $\epsilon < \infty$ and when $\epsilon = \infty$.

Case with $\epsilon < \infty$. Plugging (B14) in (6) gives w_{Lt} in function of N_t, G_t and H_t^P:

$$w_{Lt} = \begin{array}{c} \frac{\sigma-1}{\sigma}\beta^{\frac{1+\beta q}{1+q}} \left((1-\beta)\frac{1+q}{q}\right)^{\frac{(1-\beta)q}{1+q}} \frac{1}{l^{1-\beta}} \left(\frac{H_t^P}{\overline{H}}\right)^{\frac{1-\beta}{1+q}} N^{\frac{1}{\sigma-1}} \\ \times \frac{\left(G_t\left(1+\varphi w_{Lt}^{\epsilon-1}\right)^{\mu-1}+(1-G_t)\right)^{\frac{1-\beta}{1+q}}}{\left(G_t\left(\varphi w_{Lt}^{\epsilon-1}+1\right)^{\mu}+1-G_t\right)^{\frac{1}{1-\sigma}+\frac{1-\beta}{1+q}}} \end{array}, \tag{B15}$$

which replaces (A2). It is direct that when $G_\infty < 1$, we obtain (B13). In this case, we further have

$$g_\infty^{\bar{j}} = q\left(g_\infty^{w_L} - g_\infty^{w_H}\right) = -\frac{q\beta\left(\sigma-1\right)}{1+q+\beta\left(\sigma-1\right)} g_\infty^{GDP}. \tag{B16}$$

Case with $\epsilon = \infty$. In this case, (B15) becomes

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

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

Once again, following the steps of Appendix A2, we get that if $G_\infty < 1$, (B13) applies (and accordingly we also get (B16)). ■

Intuitively, as low- and high-skill wages diverge, workers switch from being low-skill to high-skill. This endogenous supply response dampens the relative decline in low-skill wages. Hence, besides securing themselves a higher future wage growth, low-skill workers who switch to a high-skill occupation also benefit the remaining low-skill workers. Since all changes in the stock of labor are driven by demand-side effects, wages and employment move in the same direction.

B4. 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 I, it is possible to interpret automation as a decline of the price of a specific equipment from infinity (the machine does not exist) to 1. Yet, our assumption that once a machine is invented, its price is constant, is crucial for deriving the general conditions under which the real wages of low-skill workers must increase asymptotically in Proposition 2. We generalize this in what follows.

Let there be two final good sectors, both perfectly competitive employing CES production technology with identical elasticity of substitution, σ. The output of sector 1, Y, is used for consumption. The output of sector 2, X, is used for machines. The two final good sectors use distinct versions of the same set of intermediate products, where we denote the use of products as $y_1(i)$ and $y_2(i)$, respectively, with $i \in [0, N]$. The two versions of product i are produced by the same supplier using production technologies that differ only in the weight on high-skill labor:

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

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

PROPOSITION B.3: *Consider three processes* $[N_t]_{t=0}^{\infty}$, $[G_t]_{t=0}^{\infty}$ *and* $[H_t^P]_{t=0}^{\infty}$ *where* $(N_t, G_t, H_t^P) \in (0,\infty) \times [0,1] \times (0,H]$ *for all* t. *Assume that* G_t, g_t^N *and* H_t^P *all admit strictly positive limits, then:*

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

$$g_{\infty}^{GDP} = \left[\psi_1 + \psi_1 \frac{\beta_1 \left(\beta_2 - \beta_1\right)}{1-\beta_2}\right] g_{\infty}^N, \tag{B17}$$

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

$$g_{\infty}^{w_L} = \frac{1}{1+\beta_1(\sigma-1)} \frac{1-\beta_2+\beta_1\left(\beta_2-\beta_1\right)\left(1-\psi_1^{-1}\right)}{1-\beta_2+\beta_1\left(\beta_2-\beta_1\right)} g_{\infty}^{GDP}. \tag{B18}$$

Proposition B.3 reduces to Proposition 2 when $\beta_2 = \beta_1$. When $\beta_2 > \beta_1$, the productivity of machine production increases faster than that of the production of Y, implying a gradual decline in the real price of machines. For given g_{∞}^N, a faster growth in the supply of machines increases the (positive) growth in the relative price of low-skill workers compared with machines, w_L/p^x, but simultaneously, it reduces the real price of machines, p^x. The combination of these two effects always implies that low-skill workers capture a lower fraction of the growth in Y. Low-skill wages are more likely to fall asymptotically for higher values of the elasticity of substitution between products, σ, as this implies a more rapid substitution away from non-automated products.

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

$$c_k\left(\alpha\left(i\right)\right) = \beta_k^{-\beta_\kappa}\left(1-\beta_k\right)^{-(1-\beta_k)}\left(w_L^{1-\epsilon}+\varphi\left(p^x\right)^{1-\epsilon}\alpha\left(i\right)\right)^{\frac{\beta_k}{1-\epsilon}} w_H^{1-\beta_k}, \tag{B19}$$

for $k \in \{1,2\}$ indexing, respectively, the production of final good and machines. As before aggregating (B19) and the price normalization gives a "productivity" condition, which replaces (6).

(B20)

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

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

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

(B21)
$$\left(G\left(w_L^{1-\epsilon}+\varphi\left(p^x\right)^{1-\epsilon}\right)^{\mu_2}+(1-G)w_L^{\beta_2(1-\sigma)}\right)^{\frac{1}{1-\sigma}} w_H^{1-\beta_2}=\frac{\sigma-1}{\sigma}\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}N^{\frac{1}{\sigma-1}}p^x.$$

Taking the ratio between these two expressions, we get

(B22)
$$\frac{\left(G\left(\left(\frac{w_L}{p^x}\right)^{1-\epsilon}+\varphi\right)^{\mu_2}+(1-G)\left(\frac{w_L}{p^x}\right)^{\beta_2(1-\sigma)}\right)^{\frac{1}{1-\sigma}} w_H^{\beta_1-\beta_2}}{\left(G\left(\left(\frac{w_L}{p^x}\right)^{1-\epsilon}+\varphi\right)^{\mu_1}+(1-G)\left(\frac{w_L}{p^x}\right)^{\beta_1(1-\sigma)}\right)^{\frac{1}{1-\sigma}}}=\frac{\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}\left(p^x\right)^{1-\beta_2+\beta_1}}{\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}}.$$

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

$$\nu_{k,x}\left(\alpha\left(i\right)\right)=\frac{\sigma-1}{\sigma}\alpha\left(i\right)\beta_k\frac{\varphi\left(p^x\right)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi\left(p^x\right)^{1-\epsilon}}, \tag{B23}$$

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

$$p^xX=NG\left(R_1(1)\nu_{1,x}(1)+R_2(1)\nu_{2,x}(1)\right). \tag{B24}$$

The zero profit condition in the machines sector gives

$$p^xX=N\left(GR_2\left(1\right)+\left(1-G\right)R_2\left(0\right)\right). \tag{B25}$$

Revenues themselves are given by

(B26)
$$R_1\left(\alpha\left(i\right)\right)=\left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1}c_1\left(\alpha\left(i\right)\right)^{1-\sigma}Y \text{ and } R_2\left(\alpha\left(i\right)\right)=\left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1}c_2\left(\alpha\left(i\right)\right)^{1-\sigma}p^xX,$$

so that (8) still holds but separately for revenues occurring from each activity and with μ_k replacing μ. Combining (8), (B23), (B24) and (B25), we get

$$\begin{array}{c}\left(G\left(1-\frac{\sigma-1}{\sigma}\beta_2\frac{\varphi(p^x)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi(p^x)^{1-\epsilon}}\right)+(1-G)\left(1+\varphi\left(\frac{w_L}{p^x}\right)^{\epsilon-1}\right)^{-\mu_2}\right)\frac{R_2(1)}{R_1(1)}\\ =G\frac{\sigma-1}{\sigma}\beta_1\frac{\varphi(p^x)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi(p^x)^{1-\epsilon}}\end{array}, \tag{B27}$$

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

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

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

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

(B28)

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

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

(B29)

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

Take the ratio between (B28) and (B29), and use (8) to obtain:

(B30)

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

Together (B20), (B22), (B27) and (B30) 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 (B22), we get

$$(p^x_t)^{1-\beta_2+\beta_1} \sim \frac{\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}}{\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}}\varphi^{\frac{\mu_2-\mu_1}{1-\sigma}} w_{Ht}^{\beta_1-\beta_2}. \tag{B31}$$

Further plugging this last relationship in (B20), we get:

$$w_{Ht} \sim \frac{\left(\frac{\sigma-1}{\sigma}\right)^{1+\frac{\beta_1}{1-\beta_2}} \varphi^{\psi_2\mu_1} \beta_1^{\beta_1} (1-\beta_1)^{1-\beta_1} \left(\beta_2^{\frac{\beta_2}{1-\beta_2}} (1-\beta_2)\right)^{\beta_1}}{G_t^{\psi_1\left(1+\frac{\beta_1(\beta_2-\beta_1)}{(1-\beta_2)}\right)} N_t^{\psi_1\left(1+\frac{\beta_1(\beta_2-\beta_1)}{(1-\beta_2)}\right)}}, \tag{B32}$$

Hence

$$g_\infty^{w_H} = \psi_1 \left(1 + \frac{\beta_1 (\beta_2 - \beta_1)}{(1-\beta_2)}\right) g_\infty^N. \tag{B33}$$

Through (B27), 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

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

In fact (B27) gives

$$\frac{R_{2,t}(1)}{R_{1,t}(1)} \sim \frac{\frac{\sigma-1}{\sigma}\beta_1}{1 - \frac{\sigma-1}{\sigma}\beta_2}. \tag{B34}$$

Using (B31) and (B32), one further gets:

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

therefore

$$g_\infty^{p_x} = -\psi_2 (\beta_2 - \beta_1) g_\infty^N < 0, \tag{B35}$$

since $\beta_2 > \beta_1$. Using that $w_{L\infty}/p_\infty^x = \infty$ and (B34) in (B30) leads to:

$$w_{Lt}\left(\frac{w_{Lt}}{p_t^x}\right)^{\epsilon-1} \sim \frac{w_{Ht}H_t^P \left(\begin{array}{c} \beta_1\left(G_t + (1-G_t)\left(\varphi\left(\frac{w_{Lt}}{p_t^x}\right)^{\epsilon-1}\right)^{1-\mu_1}\right) \\ +\beta_2\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\left(G_t + (1-G_t)\left(\varphi\left(\frac{w_{Lt}}{p_t^x}\right)^{\epsilon-1}\right)^{1-\mu_2}\right)\end{array}\right)}{\varphi G_t L\left(1-\beta_1+(1-\beta_2)\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\right)}. \tag{B36}$$

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

The reasoning then follows that of Appendix A2. If $G_\infty < 1$, then (B36) implies

$$w_{Lt}^{1+\beta_1(\sigma-1)} \sim (p_t^x)^{(\sigma-1)\beta_1} \frac{w_{Ht}H^P\beta_1(1-G_t)}{\varphi^{\mu_1}G_tL\left(1-\beta_1+(1-\beta_2)\frac{\frac{\sigma-1}{\sigma}\beta_1}{1-\frac{\sigma-1}{\sigma}\beta_2}\right)}, \tag{B37}$$

which, together with (B33) and (B35) gives (B18).

Alternatively assume that $G_\infty = 1$ and that $\lim(1-G_t)N_t^{\psi_2(1-\mu_1)\frac{\epsilon-1}{\epsilon}}$ exists and is finite. Suppose first that $\limsup(1-G_t)\left(\frac{w_{Lt}}{p_t^x}\right)^{(\epsilon-1)(1-\mu_1)} = \infty$, then there must be a sub-sequence where (B37) is satisfied, which with (B33) and (B35) leads to a contradiction with the assumption that $\lim(1-G_t)N_t^{\psi_2(1-\mu)\frac{\epsilon-1}{\epsilon}}$ exists and is finite.

If $\lim(1-G)\left(\frac{w_{Lt}}{p^x}\right)^{(\epsilon-1)(1-\mu_1)} = 0$, then (B36) gives

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

which implies with (B33) and (B35) that:

$$g_\infty^{w_L} = \frac{1}{\epsilon}\left(1-\frac{(\beta_2-\beta_1)(\epsilon-1)}{(1-\beta_2+\beta_1)}\right)g_\infty^{GDP}. \tag{B38}$$

Finally, if $\limsup(1-G_t)w_{Lt}^{(\epsilon-1)(1-\mu)}$ is finite but strictly positive, then as in Appendix A2, one can show that this requires that $\lim(1-G_t)N_t^{\frac{\psi_2}{\epsilon}(\epsilon-1)(1-\mu_1)} > 0$, from which we can derive that (B38) also holds in that case. This proves Proposition B.3 and the associated footnote in the imperfect substitutes case.

Perfect substitutes case. In the perfect substitutes case, (B20) becomes:

$$\begin{array}{c}\left(G\widetilde{\varphi}^{\beta_1(\sigma-1)}\left(p^x\right)^{\beta_1(1-\sigma)}+(1-G)w_L^{\beta_1(1-\sigma)}\right)^{\frac{1}{1-\sigma}}w_H^{1-\beta_1}\\ =\frac{\sigma-1}{\sigma}\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}N^{\frac{1}{\sigma-1}}\text{ for }w_L>p^x/\widetilde{\varphi}\end{array}, \tag{B39}$$

$$w_L^{\beta_1}w_H^{1-\beta_1}=\frac{\sigma-1}{\sigma}\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}N^{\frac{1}{\sigma-1}}\text{ for }w_L<p^x; \tag{B40}$$

(B22) becomes
(B41)

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

$$p_x=\frac{\beta_1^{\beta_1}(1-\beta_1)^{1-\beta_1}}{\beta_2^{\beta_2}(1-\beta_2)^{1-\beta_2}}w_L^{\beta_2-\beta_1}w_H^{\beta_1-\beta_2}\text{ for }w_L<p^x/\widetilde{\varphi}; \tag{B42}$$

(B27) becomes
(B43)

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

with $R_2\left(1\right)=0$ for $w_L<p^x/\widetilde{\varphi}$; and (B30) becomes
(B44)

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

$$\frac{w_LL}{w_HH^P}=\frac{\beta_1}{1-\beta_1}\text{ for }w_L<p^x/\widetilde{\varphi}. \tag{B45}$$

Together (B40), (B42) and (B45) show that we must have $w_{Lt}\geq\frac{p_t^x}{\widetilde{\varphi}}$ for t large enough, which delivers (B33) and (B35).

Assume that $G_\infty<1$, then (B44) gives (B37) from which we get that (B18) is satisfied.

Now consider the case where $G_\infty = 1$ and $\lim (1 - G_t) N_t^{\psi_2}$ exists and is finite. Then (B44) and (B43) imply

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

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

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

This proves Proposition B.3 and its associated footnote in the perfect substitutes case. ■

B5. *Intermediate steps and additional results on the baseline dynamic model*

We provide intermediate steps for the proofs of Section A3, the proof of Corollary 1 and additional analytical results on the dynamic model of Section II.

B5.1 Intermediate steps for section A3.1

In this section we derive (A14). Taking the difference between (A9) and (A10) and using (A11) we obtain:

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

Using again (A11) we get,

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

Using (A12), we can rewrite this expression as

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

Using (A8), this leads to:

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

Reordering terms and using (A13) gives (A14).

B5.2 Uniqueness of the steady state

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

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

note that H^{P*} is decreasing in g^{N*} and $\widehat{h}^{A*}$ is increasing in g^{N*}, so a sufficient condition for f to be increasing in g^{N*} is that $G^*\widehat{h}^{A*}$ is also increasing in g^{N*}. With $\widetilde{\kappa}=0$, using (A27), (A26), we get:

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

Therefore

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

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

B5.3 Proof of Lemmas of section A3.4

Proof of Lemma A.2. If $\widetilde{\kappa}=0$, $\widehat{h}_t^A$ cannot remain small forever as with positive growth in N_t, N_t and therefore w_{Lt} will become large. Since, the Poisson rate is

$\eta \left(\widehat{h}_t^A\right)^\kappa = O\left(\varphi w_{Lt}^{\epsilon-1}\right)$. This implies that G_t must start growing at a positive rate and cannot converge toward 0.

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

First assume that $G_t w_{Lt}^{\beta(\sigma-1)}$ does not tend towards 0. Then from (A32) we obtain that:

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

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

Assume now that $G_t w_{Lt}^{\beta(\sigma-1)}$ does tend towards 0. This ensures that $w_{Lt} = O\left(N_t^{\frac{1}{\sigma-1}}\right)$. Moreover, $\frac{\pi_t^A - \pi_t^N}{G_t \pi_t^A + (1-G_t)\pi_t^N} = O\left(w_{Lt}^{\beta(\sigma-1)}\right)$. Then using this in (A32), we obtain

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

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

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

Plugging this in (A8) we get:

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

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

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

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

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

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

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

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

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

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

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

Proof of Lemma A.3. Assume that G_t is bounded above 0 (note that Lemma A.2 shows that as long as $\kappa(1-\beta)+\widetilde{\kappa}<1$, it is impossible to have $G_\infty = 0$). Note that H_t^P must be bounded below otherwise there would be arbitrarily large welfare gains from increasing consumption at time t and reducing it at later time periods. As H_t^P is also bounded above (by H), then we must have (following the reasoning of Appendix A2), that $w_{Ht} = \Theta\left(N_t^{\psi}\right)$, $C_t = \Theta\left(N_t^{\psi}\right)$ and w_{Lt} is bounded below, so that $\widehat{v}_t$ and $\widehat{c}_t$ are bounded above and below and ω_t must be bounded above.

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

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

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

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

Rewriting this expression with the normalized variables and using (A20), we get:

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

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

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

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

B5.4 Behavior close to the steady-state

We now provide details on the behavior of the economy close to the steady-state. In the steady-state, using (A27) we get

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

Therefore close to the steady-state, we obtain that N_t grows at rate $g_t^N = g^{N*} + o\,(1)$, that the share of automated product obeys $G_t = G^* + o\,(1)$, with the mass of high-skill workers in automation given by $H_t^A = (1-G^*)\widehat{h}^{A*} + o\,(1)$ and the mass of high-skill workers in production given by $H_t^P = H^{P*} + o\,(1)$, with H^{P*} given by (A28). Using (A18), we obtain that $\hat{v}$ is a constant in steady-state and that wages close to the steady-state obey:

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

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

Using (A39), $\hat{\pi}^A$ is constant in steady-state and the profits made by an automated firm close to the steady-state are given by

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

(A13) then implies that the profits made by a non-automated firm π_t^N are negligible in front of π_t^A, with

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

Therefore, we get $g_\infty^{\pi^N} = g_\infty^{\pi^A} - \beta\,(\sigma-1)\,g_\infty^{w_L}$. Using (A9), the value of an automated firm is then simply given by:

$$V_t^A = \frac{\pi_t^A}{r^* - (\psi-1)\,g^{N*}} + o\left(N_t^{\psi-1}\right), \tag{B46}$$

$$\text{where } r^* = \rho + \theta\psi g^{N*} \tag{B47}$$

is the steady-state interest rate, so that $g_\infty^{V^A} = g_\infty^{\pi^A}$. Following (A10) and (A11), the normalized value of a non-automated firm obeys:

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

Therefore, one gets that for large N_t,

$$V_t^N = \frac{(1-\kappa)\,\eta G^{*\widetilde{\kappa}}\widehat{h}^{A*}}{r^* - (\psi-1)\,g^{N*} + (1-\kappa)\,\eta G^{*\widetilde{\kappa}}\widehat{h}^{A*}}V_t^A + o\left(N_t^{\psi-1}\right), \tag{B48}$$

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

B5.5 Proof of Corollary 1

Appendix B5.4 shows that in steady-state $\widehat{V}^{N*} = f\widehat{V}^{A*}$ with

$$f = \frac{(1-\kappa)\,\eta G^{*\widetilde{\kappa}}\widehat{h}^{A*}}{\rho + (\psi(\theta-1)+1)\,g_N^* + (1-\kappa)\,\eta G^{*\widetilde{\kappa}}\widehat{h}^{A*}}, \tag{B49}$$

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

Using $\widehat{V}^{N*} = \widehat{v}^*/\gamma$ and (A39), we obtain $f\frac{1}{\rho+(\psi(\theta-1)+1)g_N^*}\frac{\psi H^{P*}}{G^*} = \frac{1}{\gamma}$. Rearranging terms and using (A28), this leads to

$$g^{N*} = \frac{\frac{f}{G^*}\gamma\psi\left(H - (1-G^*)\,\widehat{h}^{A*}\right) - \rho}{\frac{f}{G^*}\psi + \frac{\theta-1}{(\sigma-1)(1-\beta)} + 1}, \tag{B50}$$

while from (A37) the growth rate when N_t is low is approximately given by

$$g^{N1} = \frac{\gamma H\psi - \rho}{\psi + \frac{\theta-1}{\sigma-1} + 1}. \tag{B51}$$

The two expressions differ by three terms: In the numerator, $H - (1-G^*)\,\widehat{h}^{A*}$ in (B50) replaces H in (B51), as some high-skill workers are hired to automate close to the steady-state, the pool of high-skill workers available for horizontal innovation or production is smaller, and this force pushes toward $g^{N*} < g^{N1}$. In the denominator, $\frac{\theta-1}{(\sigma-1)(1-\beta)}$ in (B50) replaces $\frac{\theta-1}{\sigma-1}$ in (B51) because the growth rate in the number of products has a larger impact on the economy growth rate with automation than without. This increases the effective interest rate and reduces the present value of an automated firm, therefore it also pushes toward $g^{N*} < g^{N1}$. Finally the term f/G^* in (B50) does not exist in (B51). Note that $\partial g^{N*}/(\partial f/G^*) > 0$, so that this term reflects two different forces. On one hand close to the steady-state, the value of a new firm is a fraction $f < 1$ of the value of an automated firm. On the other hand, the profits of automated firms are larger by a factor $1/G^*$ than aggregate profits, which remain a fraction $1/\sigma$of total output through the entire transitional dynamics, and this increases the value of non-automated firms. Combining (A8), (A26) and (B49), we get

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

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

B5.6 Comparative statics

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

PROPOSITION B.4: *The asymptotic growth rates of GDP g_∞^{GDP} and low-skill wages $g_\infty^{w_L}$ increase in the productivity of automation η and horizontal innovation γ.*

Therefore, in the long-run, a better automation technology (a higher η) actually benefits low-skill wages: the reason is that firms automate faster which encourages horizontal innovation.[41] During the transition, however, a higher η also means that automation takes off sooner, leading to lower low-skill wages at that point and a higher skill premium (see a numerical example in Appendix B6.3). These result preview those on the effect of taxes on automation in Section B11.6.

Proof. The proposition is established when the steady state is unique but it extends to the case of the steady states with the highest and lowest growth rates when there is multiplicity. Recall that the steady state is characterized as the solution to an equation $f\left(g^{N*}\right) = 1$ through (A30), where G^*, $\widehat{h}^{A*}$ and H^{P*} can all be written as functions of g^{N*} and parameters. Moreover, when there is a single steady state (as well as for the steady states with the highest and the lowest growth rates in case of multiplicity), f must be increasing in the neighborhood of g^{N*}.

Comparative static with respect to γ. (A27) implies that $\widehat{h}^{A*}$ is inversely proportional to γ (for given g^{N*}). Formally, we have:

$$\frac{\partial \widehat{h}^{A*}}{\partial \gamma} = -\frac{\widehat{h}^{A*}}{\gamma}. \tag{B52}$$

Differentiating (A26) and using (B52) leads to:

$$\frac{\partial G^*}{\partial \gamma} = \frac{-\kappa g^{N*} G^*}{\gamma \left(\eta \left(G^*\right)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + \left(1-\widetilde{\kappa}\right) g^{N*}\right)}, \tag{B53}$$

so that for a given g^{N*}, G^* is also decreasing in γ. Using (A28), (B52) and (B53),

[41] Corollary 1 establishes that the growth rate of the number of products is higher in a world with no automation at all than in a world with automation, but Proposition B.4 shows that conditional on automation happening ($\eta > 0$), the asymptotic growth rate in the number of products is higher when automation is easier (η is higher).

we get:

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

so that H^{P*} is increasing in γ. Note that f, defined in (A30), can be rewritten as

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

which shows that f is decreasing in γ for a given g^{N*} (H^{P*} is increasing, G^* and $\widehat{h}^{A*}$ are decreasing, and $\gamma \widehat{h}^{A*}$ is constant). Since f is increasing in g^{N*} at the equilibrium value, (A30) implies that g^{N*} increases in γ.

Comparative static with respect to η. For given g^{N*}, (A27) implies that $\widehat{h}^{A*}$ does not depend on η. Differentiating (A26), we get:

$$\text{(B54)} \qquad \frac{\partial \ln G^*}{\partial \ln \eta} = \frac{g^{N*}}{\eta (G^*)^{\widetilde{\kappa}} \left(\widehat{h}^{A*}\right)^{\kappa} + (1-\widetilde{\kappa}) g^{N*}},$$

so for given g^{N*}, G^* increases in η. (A28) implies then that

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

Using this equation together with (B54) and (A30), we obtain:

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

Using (A27), we can rewrite this as:

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

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

B6. Numerical illustration

We illustrate the results of Section II.C and further analyze the behavior of our economy through the use of numerical simulations.[42] Section B6.2 shows examples where low-skill wages temporarily decline. Section B6.3 shows the effects of changing the innovation parameter on wages and Section B6.4 gives a systematic exploration of the parameter space. Section III calibrates a richer model to the U.S. data.

B6.1 Illustrating Section II.C

Unless otherwise specified, the broad patterns described below do not depend on specific parameter choices and we simply choose "reasonable" parameters (Table B1). For convenience we loosely refer to the time period where N_t is low and the economy behaves close to a Romer model as the first phase, the period where the economy approaches its steady-state as the third phase, and the period in between as the second phase

Table B1—Baseline Parameter Specification

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

Baseline Parameters. Total stock of labor is 1 with $L = 2/3$ and $\beta = 2/3$ such that absent automation and if all high-skill workers were in production the skill premium would be 1. The initial mass of products is set low at $N_0 = 1$ to ensure we begin in Phase 1. The initial share of automated products is low, $G_0 = 0.001$, but would initially decline had we chosen a higher level. We set $\sigma = 3$ to capture an initial labor share close to 2/3. We set $\tilde{\varphi} = 0.25$ and $\epsilon = 4$, so that at $t = 0$, the profits of automated firms relative to non-automated firms are only 0.004%. The innovation parameters (γ, η, κ) are chosen such that GDP growth is close to 2% both initially and asymptotically. There is no externality from the share of automated products in the automation technology, $\widetilde{\kappa} = 0$. ρ and θ are chosen such that the interest rate is around 6% initially and asymptotically.

Figure B1 plots the evolution of the economy. Based on the behavior of automation expenditures (Panel C) we roughly delimit Phase 1 as corresponding to the first 100 years and Phase 2 as the period between year 100 and year 250.

[42] We employ the so-called "relaxation" algorithm for solving systems of discretized differential equations (Trimborn, Koch and Steger, 2008). See Appendix B7 for details.

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

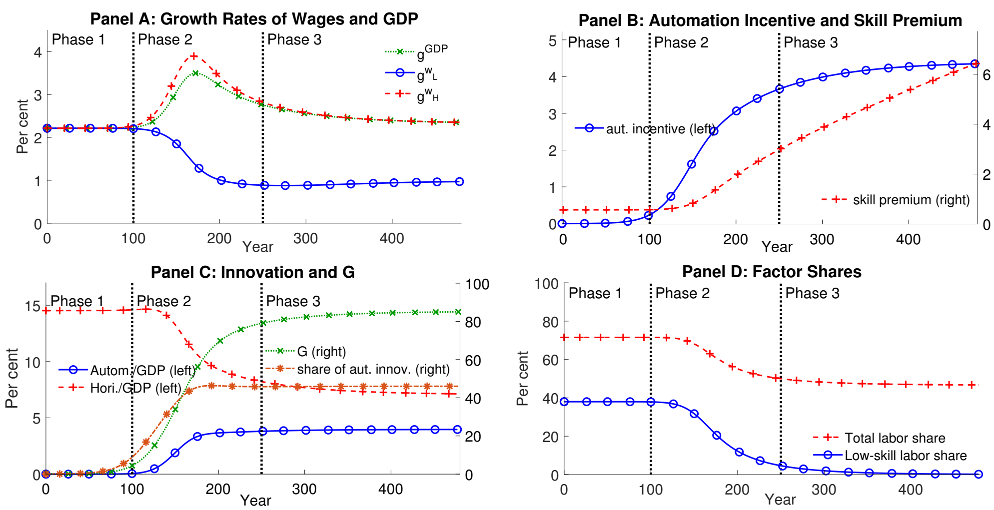

FIGURE B1. TRANSITIONAL DYNAMICS FOR BASELINE PARAMETERS.

Note: Panel A shows growth rates for GDP, low-skill wages (w_L) and high-skill wages (w_H), Panel B the incentive to automate, $\left(V_t^A - V_t^N\right)/(w_{Ht}/N_t)$, and the skill premium, Panel C the total spending on horizontal innovation and automation as well as the share of automated products (G), and Panel D the wage share of GDP for total wages and low-skill wages.

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

Wages. In the first phase, growth comes mostly from horizontal innovation and both wages grow at around 2% (Panel A). As rising low-skill wages trigger the second phase, the growth rate of high-skill wages increases to almost 4% and the growth rate of low-skill wages declines to around 1%. Though our parameter values satisfy the conditions of Proposition 1 B.ii and any increase in G_t has a negative impact on w_{Lt}, the growth in N_t is sufficient to ensure that low-skill wages grow at a positive rate throughout (see Section B6.2 for counter-examples). Finally, in the third phase, the growth rate of low-skill wages stabilizes at around 1% and the skill premium keeps rising but more slowly than previously.

Factor shares. Panel D of Figure B1 plots the labor share and the low-skill labor share. With machines as intermediate inputs, capital income corresponds to aggregate profits, which are a constant share of output. High-skill labor in production also earns a constant share of output. Both correspond to a rising share of GDP in Phase 2 as during this time period, the ratio Y/GDP increases since machines expenditures are excluded from GDP. The low-skill labor share is nearly constant in Phase 1 but declines with automation in Phase 2 and approaches 0 in Phase 3. The total labor share of GDP follows a similar pattern—but its decline is less marked since the high-skill share increases. This occurs despite an increase in the share of high-skill workers in innovation which raises the labor share (see (9)). Yet, because of this effect, the drop in the labor share can be delayed relative to the rise in the skill premium for some parameter values.

Wealth and consumption. Figure B2 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 B1.A), which is in line with a stable ratio of total R&D expenses over GDP across the three phases (Figure B1.D). In the absence of any financial constraints, low-skill and high-skill consumption must grow at the same rate, with high-skill workers consuming more since they have a higher income (Panel B). Since low-skill labor income becomes a negligible share of GDP, while the high-skill labor share increases, a constant consumption ratio can only be achieved if high-skill workers borrow from low-skill workers in the long-run. This is illustrated in Panel C, which shows the share of assets held by low-skill workers, under the assumption that initially assets holdings per capita are identical for low-skill and high-skill workers (so that low-skill workers hold 2/3 of the assets in year 0, since with these parameters $H/L = 1/2$). Initially, low-skill and high-skill income grow at a constant rate so that the share of assets held by low-skill workers is stable; but, in anticipation of a lower growth rate for low-skill wages than for high-skill wages, low-skill workers start saving more, and the share of assets they hold increases. This share eventually reaches more than 100%, meaning that the high-skill workers net worth becomes negative. Panel D shows that since profits become a higher share of GDP (an effect which dominates a temporary increase in the interest rate in Phase 2), the wealth to GDP

ratio increases in phase 2, such that its steady state value is nearly 3 times higher than its original value.

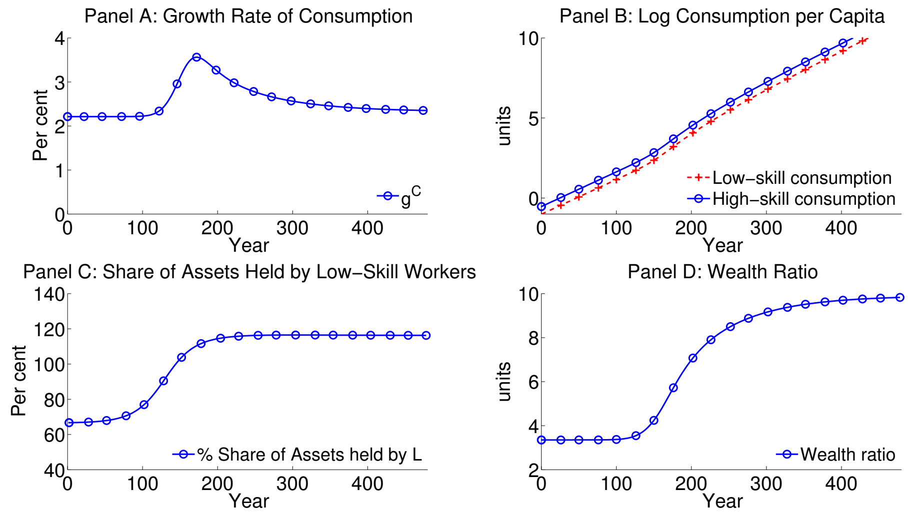

Figure B2. Consumption and wealth for baseline parameters.

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

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

Growth decomposition. Figure B3 performs a growth decomposition exercise for low-skill and high-skill wages by separately computing the instantaneous contribution of each type of innovation. We do so by performing the following thought experiment: at a given instant t, for given allocation of factors, suppose that all innovations of a given type fail. By how much would the growth rates of w_L and w_H change?[43] In Phase 1, there is little automation, so wage growth

[43] More specifically we can write $w_{Lt} = f(N_t, G_t, H_t^P)$, using equations (7) and (6). Differentiating with respect to time and using equation (A8) gives:

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

Figure B3 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

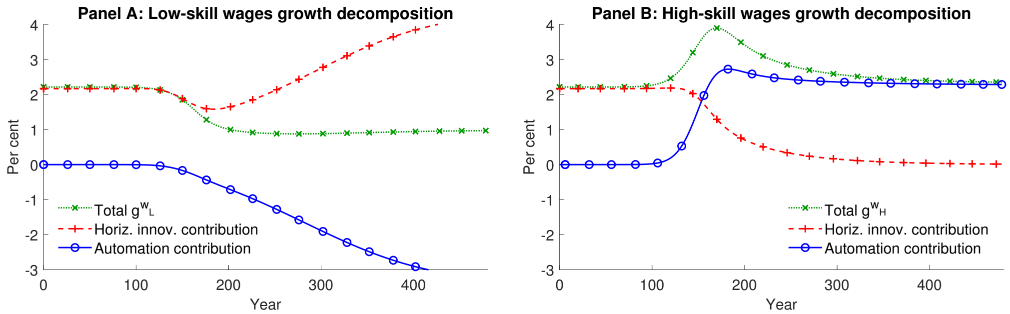

FIGURE B3. GROWTH DECOMPOSITION.

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.

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 II.D, an increase in the effectiveness of the automation technology, η, will have positive long-term consequences. A decomposition of g_t^{GDP} would look similar to the decomposition of $g_t^{w_H}$: while instantaneous growth is initially almost entirely driven by horizontal innovation, automation becomes increasingly important in explaining it (long-run growth, however, is ultimately determined by the endogenous rate of horizontal innovation).

B6.2 Negative growth for low-skill wages

This section presents two examples with negative growth for low-skill wages. We ensure temporary negative growth in low-skill wages in Figure B4 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 $\left(V_t^A - V_t^N\right) / \left(w_{Ht}/N_t\right)$ has already significantly risen (Panel B). Yet, Phase 2 is more intense once it gets started, partly because of the

for w_{Ht}.

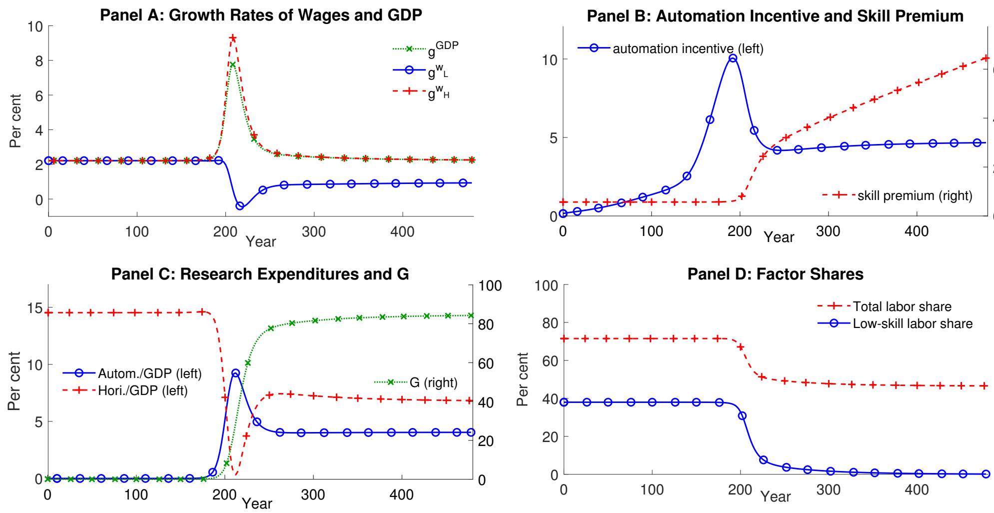

Figure B4. Transitional Dynamics with temporary decline in low-skill wages with an automation externality.

Note: same as for Figure B1 but with an automation externality of $\tilde{\kappa} = 0.49$.

sharp increase in the productivity of the automation technology (following the increase in G_t) and partly because low-skill wages are higher. Intense automation puts downward pressure on low-skill wages. At the same time, horizontal innovation drops considerably, both because new firms are less competitive than their automated counterparts, and because the high demand for high-skill workers in automation innovations increases the cost of inventing a new product. This results in a short-lived decline in low-skill wages. Indeed, the decline in w_{Lt} (and increase in high-skill wage w_{Ht}) lowers the incentive to automate (Panel B), which in return reduces automation. Note that in a model with endogenous adaptation where automation involves the payment of a fixed cost every period instead of a R&D sunk cost, it would not be possible to obtain even a temporary decline in low-skill wages as firms would stop paying the fixed cost as soon as wages decline. Here, the discounted value of the difference in profits between automated and non-automated firms stays at a high level throughout ensuring that automation nonetheless remains at a high level.

Low-skill wages can also drop for $\tilde{\kappa} = 0$ as shown in Figure B5 where low-skill wages slightly decline for a short time period. The associated parameters are given in Table B2. The crucial parameter change is an increase in κ, such that the automation technology is less concave. This delays Phase 2, which is then more intense and leads to a sharp increase in high-skill wages, reducing considerably horizontal innovation.

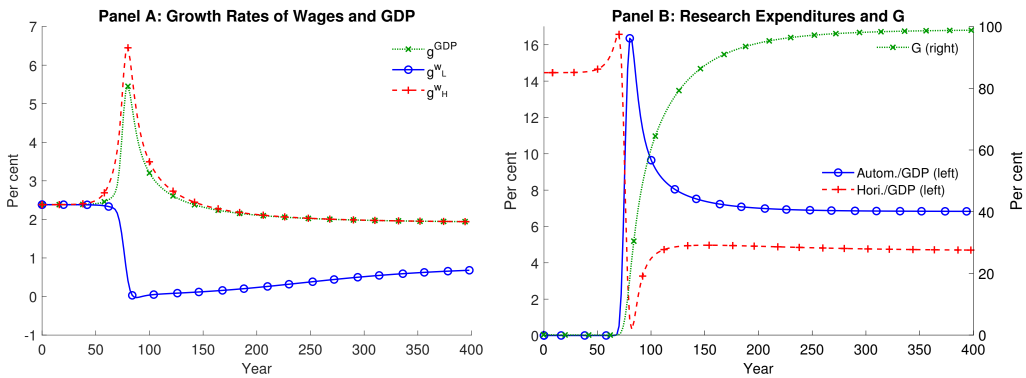

FIGURE B5. TRANSITIONAL DYNAMICS WITH TEMPORARY DECLINE IN LOW-SKILL WAGES WITHOUT AN AUTOMATION EXTERNALITY.

TABLE B2—BASELINE PARAMETER SPECIFICATION

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

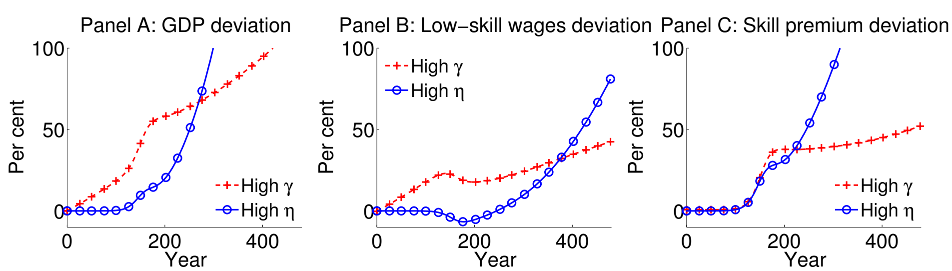

FIGURE B6. DEVIATIONS FROM BASELINE MODEL FOR MORE PRODUCTIVE HORIZONTAL INNOVATION TECHNOLOGY (γ) AND MORE PRODUCTIVE AUTOMATION TECHNOLOGY (η).

B6.3 The effect of the innovation parameters

Figure B6 shows the impact (relative to the baseline case) of increasing productivity in the automation technology to $\eta = 0.4$ (from 0.2) and the productivity in the horizontal innovation technology to $\gamma = 0.32$ (instead of 0.3). A higher η initially has no impact during Phase 1, but it moves Phase 2 forward as investing in automation technology is profitable for lower level of low-skill wages. Since automation occurs sooner, the absolute level of low-skill wages drops relative to the baseline case (Panel B), which leads to a fast increase in the skill premium. However, as a higher η means that new firms automate faster, it encourages further horizontal innovation. A faster rate of horizontal innovation implies that the skill premium keeps increasing relative to the baseline, but also that low-skill wages are eventually larger than in the baseline case. A higher productivity for horizontal innovation, γ, implies that GDP and low-skill wages initially grow faster than in the baseline. Therefore Phase 2 starts sooner, which explains why the skill premium jumps relative to the baseline case before increasing smoothly.

B6.4 Systematic comparative statics

In this section we carry a systematic comparative exercise with respect to the parameters of the model, namely

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

Figures B7.A,B,C show that a higher elasticity of substitution across products σ reduces the growth rate of the economy (the elasticity of output with respect to the number of products is lower), which leads to a delayed transition. The asymptotic growth rate of low-skill wages is a smaller fraction of that of high-skill wages (following Proposition 2), since automated products are a better substitute for non-automated ones. Figures B7.D,E,F show that the elasticity of substitution between machines and low-skill workers in automated firms, ϵ, plays a limited role (as long as $\mu < 1$), a higher elasticity reduces the growth of low-skill wages and increases that of high-skill wages during Phase 2. Figures B7.G,H,I show that a lower factor share in production for high-skill workers (a higher β) increases the growth rate of the economy (high-skill wages are lower which favors innovation). As a result, Phase 2 occurs sooner. Besides, following Proposition 2, the asymptotic growth rate of low-skill wages is a lower fraction of that of high-skill wages (the cost advantage of automated firms being larger).

Figures B8.A,B,C show that a higher discount rate ρ reduces the growth rate of the economy, which slightly postpones Phase 2. At the time of Phase 2, the growth rate of low-skill wages is not affected much by the discount rate: on one hand, since low-skill wages are lower Phase 2 is postponed, which favor low-skill wages'

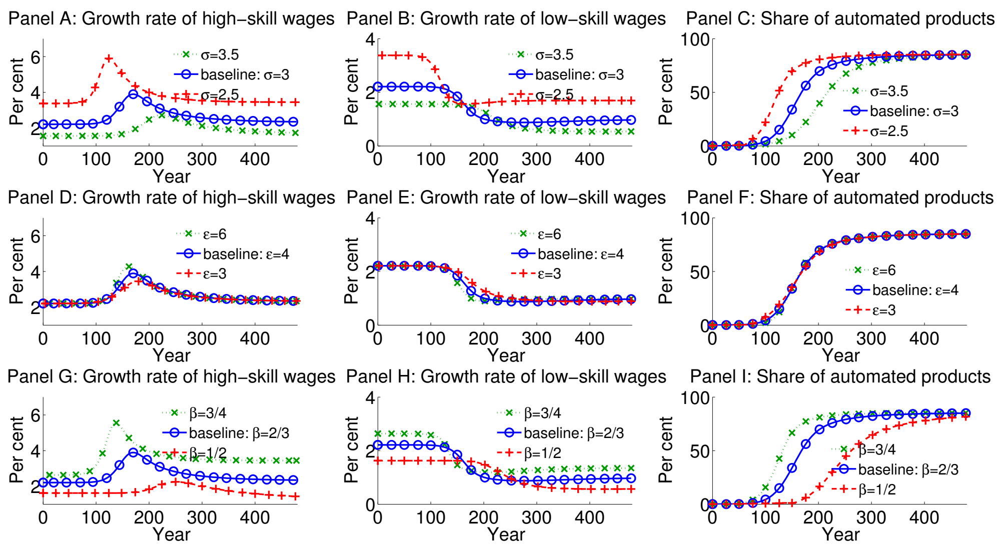

FIGURE B7. COMPARATIVE STATICS WITH RESPECT TO THE ELASTICITY OF SUBSTITUTION ACROSS PRODUCTS (σ), THE ELASTICITY OF SUBSTITUTION BETWEEN MACHINES AND LOW-SKILL WORKERS IN AUTOMATED FIRMS (ϵ) AND THE FACTOR SHARE OF LOW-SKILL WORKERS AND MACHINES IN PRODUCTION (β).

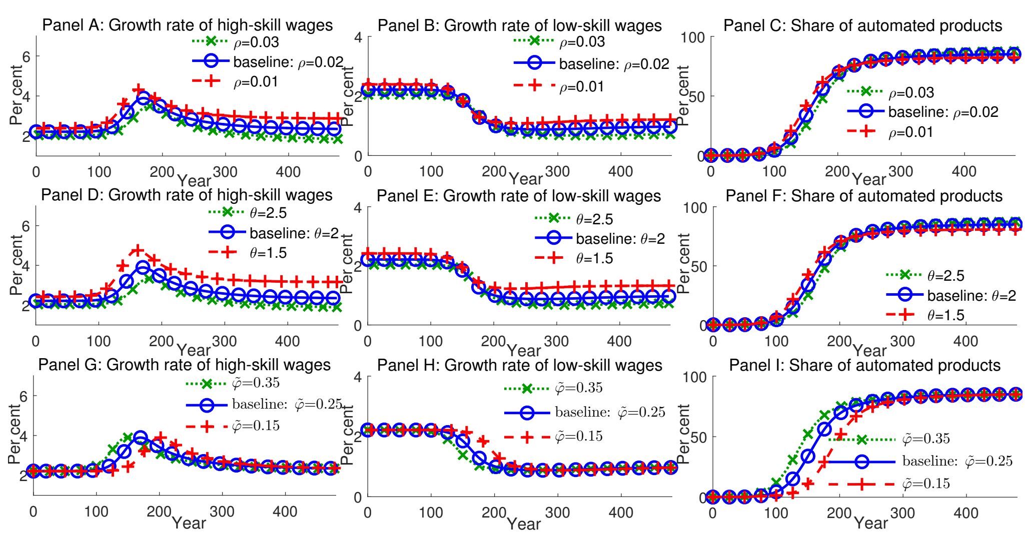

FIGURE B8. COMPARATIVE STATICS WITH RESPECT TO THE DISCOUNT RATE (ρ), THE INVERSE ELASTICITY OF INTERTEMPORAL SUBSTITUTION (θ) AND THE PRODUCTIVITY OF MACHINES ($\tilde{\varphi}$)

growth, but on the other hand, horizontal innovation is lower which negatively affects low-skill wages. A lower elasticity of intertemporal substitution (a higher θ) has a similar effect on the economy's growth rate (Figures B8.D,E,F). Figures B8.G,H,I show that the productivity of machines ($\widetilde{\varphi}$) only affects the timing of Phase 2 (Phase 2 occurs sooner when machines are more productive).

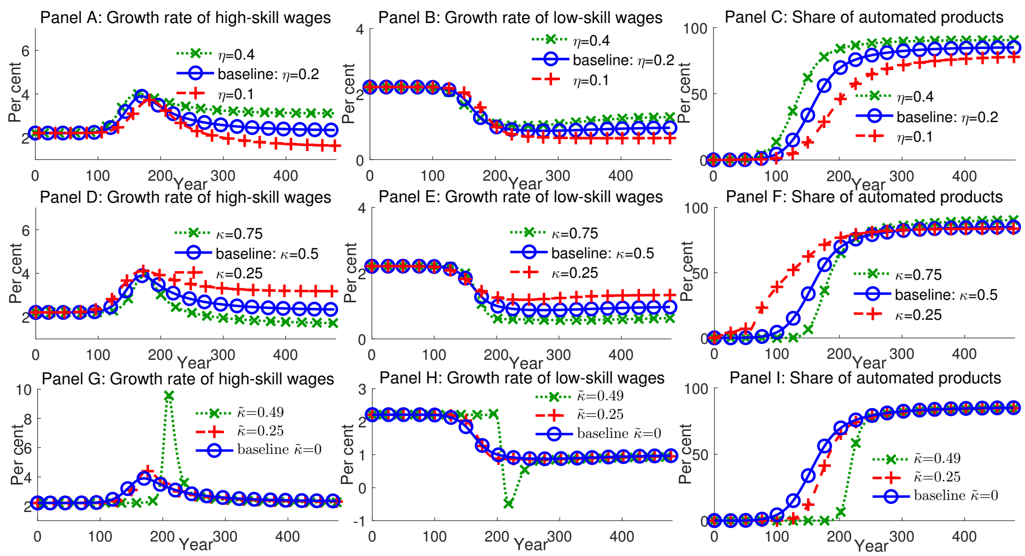

FIGURE B9. COMPARATIVE STATICS WITH RESPECT TO THE AUTOMATION PRODUCTIVITY (η), THE CONCAVITY OF THE AUTOMATION TECHNOLOGY (κ) AND THE AUTOMATION EXTERNALITY ($\tilde{\kappa}$)

The comparative statics with respect to the automation technology shown in Figures B9.A,B,C follow the pattern described in the text. A less concave automation technology (higherκ) delays Phase 2 and reduces the economy's growth rate. It particularly affects the growth rate of low-skill wages in Phase 2 (as the increase in automation expenses comes more at the expense of horizontal innovation)—see Figures B9.D,E,F. The role of the automation externality has already been discussed in the text, Figures B9.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 B10.A,B,C show the impact of the horizontal innovation parameter γ, which was already discussed in the text. Figures B10.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 B11.A,B,C show that a higher initial number of products simply advance the entire evolution of the economy. Figures B11.D,E,F show that a higher initial value for the share of automated products (even as high as the

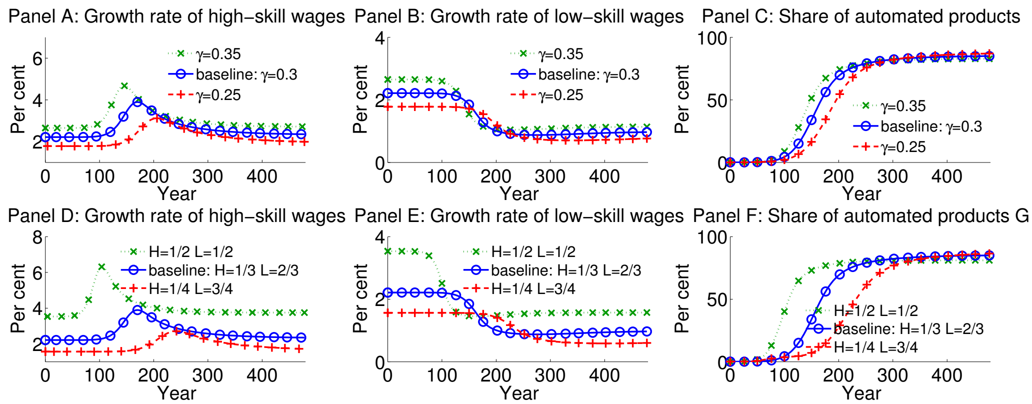

Figure B10. Comparative statics with respect to the horizontal innovation productivity (η) and the skill ratio (H/L)

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.

B7. *Simulation technique*

In the following we describe the simulation techniques employed in Appendix B6 for the baseline model presented in I. The approach for the extensions and the quantitative exercise of Section III follow straightforwardly. Let $\mathbf{x}_t \equiv (n_t, G_t, \hat{h}_t^A, \chi_t, \omega_t)$ and note that equation (A24) defines ω implicitly. We can therefore write equations (A7), (A8), (A14) and (A15) as a system of autonomous differential equations $(\dot{n}_t, \dot{G}_t, \dot{\hat{h}}_t^A, \dot{\chi}_t) = F(\mathbf{x}_t)$ with initial conditions on state variables as (n_0, G_0) and an auxiliary equation of $\omega_t = \vartheta(\mathbf{x}_t)$.

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

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

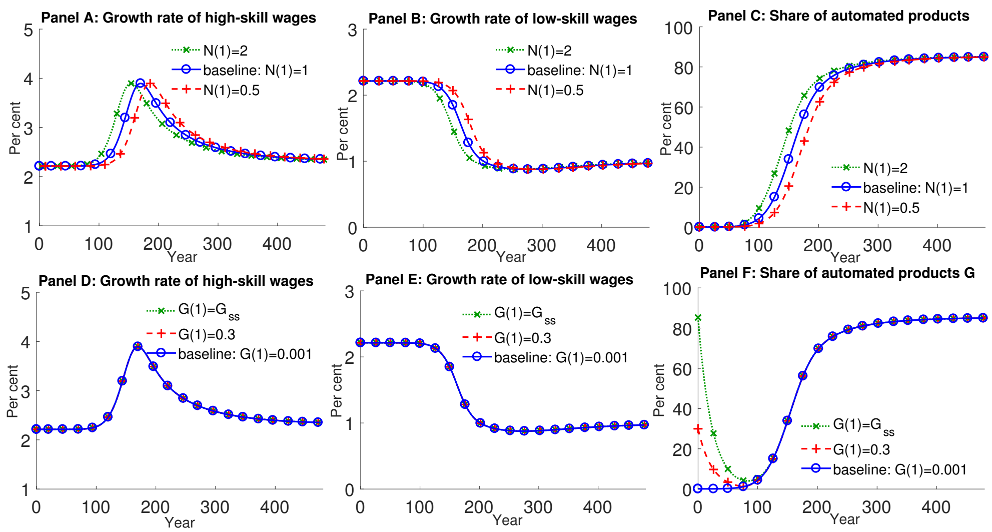

Figure B11. Comparative statics with respect to the initial number of products N_0 and the initial share of automated products G_0

with T corresponding errors for ω_t :

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

Following Lemma A.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 χ_T accordingly. Together with the initial conditions ($n_1 = n^{start}$ and $G_1 = G^{start}$) this leads to a vector of errors:

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

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

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

for a $5T \times 5T$ diagonal weighting matrix, $\mathbf{W}$, and the $5T$ vector $\mathbf{x}$. For $dt \to 0$ and $T \to \infty$ $\mathbf{S}(\mathbf{x})'\mathbf{W}\mathbf{S}(\mathbf{x}) \to 0$. For the simulations we set $dt = 2$ and $T = 2000$. We accept the solution when $\mathbf{S}(\hat{\mathbf{x}})'\mathbf{W}\mathbf{S}(\hat{x}) < 10^{(-7)}$, but the value is typically less than $10^{(-10)}$. The choice of weighting matrix matters somewhat for the speed of convergence, but is inconsequential for the final result. With the solution $\{\hat{\mathbf{x}}_t, \hat{\omega}_t\}_{t=1}^T$ in hand, it is straightforward to find remaining predicted values.

B8. 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.

B8.1 Characterizing the optimal allocation

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

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

such that

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

with

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

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

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

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

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

$$H_t^D \geq 0.$$

The first order condition with respect to C_t gives

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

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

The first order conditions with respect to X_t implies that

$$\frac{\partial F}{\partial X_t} = 1, \tag{B55}$$

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

$$w_{Lt} = \frac{\partial F}{\partial L_t^A} = \frac{\partial F}{\partial L_t^N} \text{ and } w_{Ht} = \frac{\partial F}{\partial H_t^{P,A}} = \frac{\partial F}{\partial H_t^{P,N}}, \tag{B56}$$

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

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

$$\rho\widetilde{\zeta}_t - \dot{\widetilde{\zeta}}_t = \begin{array}{c} \widetilde{\lambda}_t \frac{\partial F}{\partial N_t^N} + \widetilde{\zeta}_t \gamma H_t^D + \left(\widetilde{\xi}_t - \widetilde{\zeta}_t\right) \eta \left(H_t^A\right)^{\kappa} \left(N_t^N\right)^{-\kappa} \\ \times \left(N_t^A\right)^{\widetilde{\kappa}} \left(\left(1-\widetilde{\kappa}\right) N_t^N + \left(1-\kappa\right) N_t^A\right) \left(N_t^N + N_t^A\right)^{\kappa-\widetilde{\kappa}-1} \end{array}, \tag{B57}$$

$$\rho\widetilde{\xi}_t - \dot{\widetilde{\xi}}_t = \begin{array}{c} \widetilde{\lambda}_t \frac{\partial F}{\partial N_t^A} + \widetilde{\zeta}_t \gamma H_t^D + \left(\widetilde{\xi}_t - \widetilde{\zeta}_t\right) \eta \left(H_t^A\right)^{\kappa} \\ \times \left(N_t^N\right)^{1-\kappa} \left(N_t^A\right)^{\widetilde{\kappa}-1} \left(\widetilde{\kappa} N_t^N + \kappa N_t^A\right) \left(N_t^N + N_t^A\right)^{\kappa-\widetilde{\kappa}-1} \end{array}. \tag{B58}$$

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

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

Further the (shadow) interest rate is given by $r_t = \rho + \theta \frac{\dot{C}_t}{C_t} = \rho - \frac{\dot{\lambda}_t}{\lambda_t}$. Using that

$H_t^A = (1-G_t)\, N_t h_t^A$, we can rewrite (B57) and (B58) as:
(B59)

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

(B60) $$r_t \xi_t = \pi_t^A + \zeta_t g_t^N + (\xi_t - \zeta_t)\, \eta\, (G_t)^{\widetilde{\kappa}}\, N_t^{\kappa} \left(h_t^A\right)^{\kappa} (1-G_t) \left(\widetilde{\kappa}\frac{1-G_t}{G_t} + \kappa\right) + \dot{\xi}_t.$$

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

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

(B61) $$w_{Ht} \geq \zeta_t \gamma N_t,$$

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

(B62) $$w_{Ht} = (\xi_t - \zeta_t)\, \kappa \eta\, (G_t)^{\widetilde{\kappa}}\, N_t^{\kappa} \left(h_t^A\right)^{\kappa-1}.$$

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

B8.2 System of differential equations and steady state

After having introduced the same variables as in the equilibrium case, one can follow the same steps and derive a system of differential equation in $\left(n_t, G_t, \widehat{h}_t^A, \chi_t\right)$ which characterizes the solution (when there is positive growth). Equations (A7) and (A8) still hold, while equations (A14) and (A15) are replaced with
(B63)

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

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

g_t^N is still given by (A23), $\frac{\widehat{\pi}_t^A}{\widehat{v}_t}$, H_t^P and ω_t are now given by

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

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

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

which replace (A20), (A22) and (A24).

One can then solve for a steady state of this system with $G^* > 0$ (and $\left(g^N\right)^* > 0$ so that $n^* = 0$). (A26) and (A28) still apply, but (A27) is replaced with

(B64) $$\widehat{h}^{A*} = \frac{\kappa}{\gamma}\frac{\rho+(\theta-1)\,\psi g^{N*}}{1-\kappa+(1-G^*)\,(\kappa-\widetilde{\kappa})},$$

and (A30) with

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

which is obtained by fixing $\dot{\widehat{h}}_t^{A} = 0$ in (B63) using (A26) and (B64). For g^{N*} large enough (but finite—and, in particular smaller than γH), H^{P*} is arbitrarily small, while for the same value G^* and $\widehat{h}^{A*}$ are bounded below and above. As before, this establishes that for g^{N*} large enough $f^{sp}\left(g^{N*}\right) > 1$. Furthermore $f^{sp}(0) = f(0)$, therefore condition A25 is also a sufficient condition for the existence of a steady state with positive growth and $G^* > 0$ for the system of differential equations.

B8.3 Decentralizing the optimal allocation

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

$$r_t V_t^N = \pi_t^N - \Omega_t^A w_{Ht} h_t + \eta\left(G_t\right)^{\widetilde{\kappa}} N_t^{\kappa}\left(h_t^A\right)^{\kappa}\left(V_t^A - V_t^N\right) + \dot{V}_t^N, \tag{B65}$$

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

$$\kappa\eta\left(G_t\right)^{\widetilde{\kappa}} N_t^{\kappa}\left(h_t^A\right)^{\kappa-1}\left(V_t^A - V_t^L\right) = \Omega_t^A w_{Ht}, \tag{B66}$$

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

$$\gamma N_t V_t^N = \Omega_t^H w_{Ht}, \tag{B67}$$

instead of (17). For Ω_t^A and Ω_t^H to decentralize the optimal allocation it must be that these 4 equations hold together with (B59), (B60), (B61) and (B62).

Using (B61) and (B67), we then get that Ω_t^H must satisfy

$$\Omega_t^H \zeta_t = V_t^N, \tag{B68}$$

similarly, using (B62) and (B66), we get

$$\Omega_t^A (\xi_t - \zeta_t) = V_t^A - V_t^L. \tag{B69}$$

Plugging (B68) and (B69) in (B65), we get that

$$r_t \zeta_t = \frac{\pi_t^N}{\Omega_t^H} - \frac{\Omega_t^A}{\Omega_t^H} w_{Ht} h_t + \eta (G_t)^{\widetilde{\kappa}} N_t^\kappa \left(h_t^A\right)^\kappa \frac{\Omega_t^A}{\Omega_t^H} (\xi_t - \zeta_t) + \frac{\dot{\Omega}_t^H}{\Omega_t^H} \zeta_t + \dot{\zeta}_t. \tag{B70}$$

Similarly, using (B69) and the difference between (14) and (B65) gives:
(B71)

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

Combining (B70) with (B59), using (B62) and (B61) and the definition of Ω_t^A and Ω_t^H, we get:

$$\dot{s}_t^H = \begin{array}{c} \frac{\gamma \widehat{\pi}_t^N}{\widehat{v}_t} s_t^H - \left(1 - s_t^H\right) g_t^N \\ + \frac{\gamma \widehat{h}_t^A}{\kappa} \left(\left(1 - s_t^A\right)(1 - \kappa) + \left(1 - s_t^H\right)\left(\widetilde{\kappa}(1 - G_t) + \kappa G_t - 1\right)\right) \end{array}. \tag{B72}$$

Similarly combining (B71) with the difference between (B60) and (B59) and using (B61) gives:

$$\dot{s}_t^A \frac{\left(\widehat{h}_t^A\right)^{1-\kappa}}{\eta G_t^{\widetilde{\kappa}}} = \kappa \frac{\widehat{\pi}_t^A - \widehat{\pi}_t^N}{\widehat{v}_t} s_t^A - \widetilde{\kappa} \left(1 - s_t^A\right) \widehat{h}_t^A \frac{1 - G_t}{G_t}. \tag{B73}$$

Therefore, in steady state, we have

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

Note from (B73) that the share of automated products, s_t^A, must always be non-negative, otherwise it cannot converge to a positive value, therefore $s_t^A \geq 0$ everywhere (and in fact > 0 if $\widetilde{\kappa} \neq 0$). Furthermore, if $\widetilde{\kappa} = 0$, $s_t^A = 0$ everywhere, the only externality in automation comes from the total number of products, therefore the equilibrium features the optimal amount of automation investment (when the monopoly distortion is corrected and the optimal subsidy to horizontal innovation is implemented).

(B72) gives the steady state value of the subsidy to horizontal innovation as:

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

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

B8.4 Transitional dynamics for the social planner case
Figure B12 plots the transitional dynamics for the optimal allocation in our baseline case (with $\widetilde{\kappa} = 0$) and in the case where $\widetilde{\kappa} > 0$ analyzed in Figure B4. 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 B1.A and Figure B4.A, the overall dynamics look quite similar but the growth rates are higher in the social planner case, and the transition to phase 2 now happens roughly at the same time with and without the automation externality, while in the equilibrium it is considerably delayed in the presence of the externality (as, effectively, the productivity of the automation technology is initially very low). In both cases, the social planner maintains a positive subsidy to horizontal innovation. When $\widetilde{\kappa} = 0$ (without the automation externality), the subsidy to automation is 0, while when $\widetilde{\kappa} > 0$ there is a positive subsidy to automation, which is the largest in Phase 1. This subsidy explains why Phase 2 now starts at around the same time.

B9. An endogenous supply response in the skill distribution: dynamic model

In this Appendix we revisit the model with endogenous skill supply of Section B3 and we now characterize its behavior with endogenous innovation. We can derive the dynamic system as in the baseline model (see details below in Appendix B9.1). Lemma A.1 can then be extended and in fact the steady state values $(G^*, \hat{h}^{A*}, g^{N*}, \chi^*)$ are the same as in the model with a fixed high-skill labor supply $\bar{H}$.

Figure B13 shows the transitional dynamics for this model when the common parameters are the same as in Table B1, $\overline{H} = 1/3$ (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 B1, but the gap in steady-state between the low-skill growth rate and the high-skill growth rate is a bit smaller. In addition Panel B shows that the skill ratio increases from Phase 2 and Panel A shows that the growth rate is lower in Phase 1 as the mass of high-skill workers is lower then.

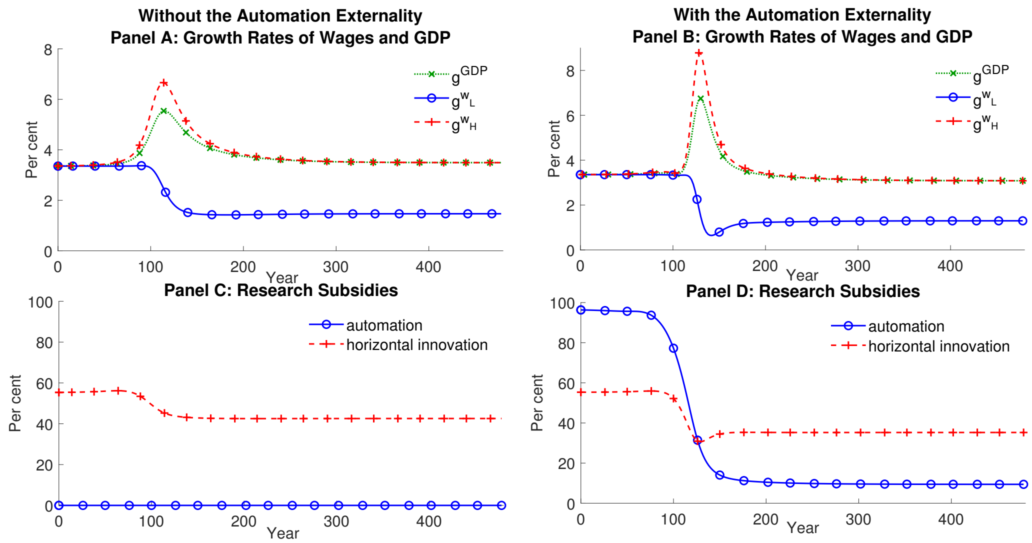

FIGURE B12. TRANSITIONAL DYNAMICS IN THE SOCIAL PLANNER CASE. PANEL A AND B, BASELINE CASE. PANEL C AND D, WITH $\tilde{\kappa} = 0.5$

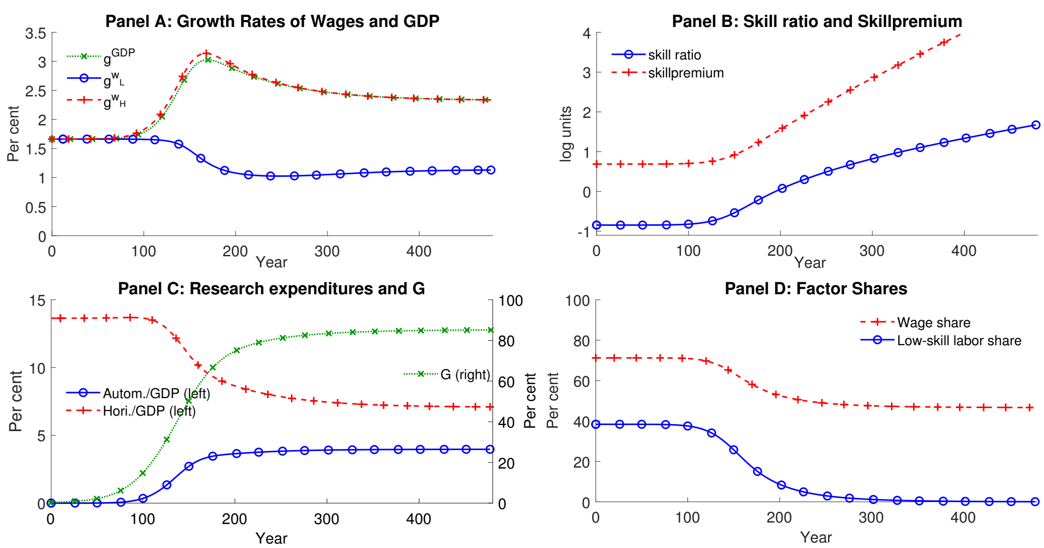

FIGURE B13. TRANSITIONAL DYNAMICS FOR MODEL WITH ENDOGENOUS SKILL SUPPLY.

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

B9.1 Details on the dynamic system

It is convenient to redefine $n_t \equiv N_t^{\frac{-\beta}{(1-\beta)}\frac{(1+q)}{1+q+\beta(\sigma-1)}}$, we can then write the entire dynamic system as a system of differential equations in $\left(n_t, G_t, \widehat{h}_t^A, \chi_t\right)$ with two auxiliary variables ω_t and $\hat{j}_t \equiv \overline{j}_t n_t^{-\frac{q}{1+q}}$. Equations (A7) is now given by

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

(A8), (A14), (A15), (A20), (A22) still apply and equation (A23) as well provided that H is replaced by H_t given by (B11). ω_t is implicitly defined by:

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

which replaces (A24) and is a rewriting of (B15) and $\hat{j}_t$ is given by

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

which is derived using (B12) and (B14).

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

B10. Description of the Quantitative Model and Analytical Results

B10.1 Set-up

To avoid repetitions, we already include the taxes of Section B11.6, namely, we assume that there is a tax τ_m on the rental rate of equipment and a tax τ_a on high-skill workers in automation innovations. The solution follows similar steps to the baseline case. We denote by $\widetilde{r}_t$ the gross rental rate of machines and by Δ their depreciation rate, such that:

(B74) $$\widetilde{r}_t = r_t + \Delta.$$

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

$$c\left(w_L, w_H, \widetilde{r}, \alpha\left(i\right)\right) = \frac{\left(w_L^{1-\epsilon} + \alpha(i)\varphi\left(\widetilde{r}^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}\right)^{\frac{\beta_1}{1-\epsilon}} w_H^{\beta_2} \widetilde{r}^{\beta_3}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}} \tag{B75}$$

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

$$\frac{\sigma N^{\frac{1}{1-\sigma}}}{\sigma-1}\frac{w_H^{\beta_2}\widetilde{r}^{\beta_3}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}\left(G\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} + \left(1-G\right)w_L^{\beta_1(1-\sigma)}\right)^{\frac{1}{1-\sigma}} = 1. \tag{B76}$$

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

$$\frac{w_H H^P}{w_L L} = \frac{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}\right)\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} + \beta_2\left(1-G\right)w_L^{\beta_1(1-\sigma)}}{\beta_1\left(G w_L^{1-\epsilon}\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu-1} + \left(1-G\right)w_L^{\beta_1(1-\sigma)}\right)}. \tag{B77}$$

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

$$\frac{\widetilde{r}K}{w_H H^P} = \frac{\left(\begin{array}{c} G\left(\beta_3 + \frac{\beta_1(1-\beta_4)\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{(1+\tau_m)\left(w_L^{1-\epsilon}+\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}\right)}\right)\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} \\ +\beta_3\left(1-G\right)w_L^{\beta_1(1-\sigma)} \end{array}\right)}{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}{w_L^{1-\epsilon}+\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon}}\right)\left(\varphi\left(\left(\left(1+\tau_m\right)\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu} + \beta_2\left(1-G\right)w_L^{\beta_1(1-\sigma)}}. \tag{B78}$$

B10.2 Effect of technology on wages

First note that one can rewrite (B77)

$$\frac{w_H H^P}{w_L L} = \frac{G\left(\beta_2 + \beta_1\beta_4\frac{\Phi}{1+\Phi}\right)\left(\Phi+1\right)^{\mu} + \beta_2\left(1-G\right)}{\beta_1\left(G\left(\Phi+1\right)^{\mu-1} + \left(1-G\right)\right)}, \tag{B79}$$

where we defined
(B80)

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

In (B79), the RHS is increasing in w_L and decreasing in w_H/w_L for given G, $\widetilde{r}$. Therefore, this equation defines the relative demand curve in the w_L, w_H space as rotating counter-clockwise (when $G > 0$) when w_L increases. Plugging (B79) in (B76) then defines w_L uniquely as a function of N, G, $\widetilde{r}$ and H^P. We can then derive the effect of changes in G and N for given H^P and $\widetilde{r}$ (i.e. when K is perfectly elastically supplied) on wages, the skill premium and the labor share as follows:

PROPOSITION B.5: *Consider the equilibrium (w_L, w_H) determined by equations (B79) and (B76). 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 (B76) as:
(B81)

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

Using that (B79) establishes $\frac{w_H}{w_L}$ as an increasing function of w_L otherwise independent of N, we get that (B81) implies that w_L and therefore w_H/w_L (when $G > 0$) and w_H itself must increase in N.

(B79) 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 (B76) the direct effect of an increase in G is to decrease the LHS (because $\left(\varphi \left(\left((1+\tau_m)\,\widetilde{r}\right)^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right)^{\mu} > w_L^{\beta_1(1-\sigma)}$), in addition an increase in G would reduce w_L which further reduces the LHS. To maintain the inequality, it must be that w_H increases. Therefore in this case too, w_H increases in G and so does w_H/w_L.

This model is isomorphic to the previous one when $\beta_4 = \beta_3 = 0$ (with $\varphi\left((1+\tau_m)\,\widetilde{r}\right)^{1-\epsilon}$

replacing φ), therefore the impact of a change of G on w_L is also ambiguous.

The labor share is now given by

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

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

$$Y = \frac{\sigma}{\sigma - 1} \left(w_L L + w_H H^P + \widetilde{r} K + T_m \right), \tag{B82}$$

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

$$\frac{T_m}{w_H H^P} = \frac{G \frac{\tau_m \beta_1 (1-\beta_4) \varphi \left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon}}{(1+\tau_m)\left(w_L^{1-\epsilon} + \varphi \left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} \right)} \left(\varphi \left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right)^{\mu}}{G \left(\beta_2 + \frac{\beta_1 \beta_4 \varphi \left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon}}{w_L^{1-\epsilon} + \varphi \left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon}} \right) \left(\varphi \left(((1+\tau_m)\widetilde{r})^{1-\beta_4} w_H^{\beta_4} \right)^{1-\epsilon} + w_L^{1-\epsilon} \right)^{\mu} + \beta_2 (1-G) w_L^{\beta_1(1-\sigma)}}. \tag{B83}$$

Then, we obtain:

$$LS = \frac{w_L L + w_H H}{\frac{\sigma}{\sigma-1} (w_L L + w_H H^P + \widetilde{r} K + T_m) + (1+\tau_a) w_H (H - H^P)}. \tag{B84}$$

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

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

Using (B77), (B78), (B83) and (B80), we obtain:

$$\frac{\widetilde{r} K + T_m}{w_L L + w_H H} = \frac{\beta_3}{\beta_2 + \beta_1} + \frac{\frac{(1-\beta_4)\beta_1}{\beta_2+\beta_1} G \frac{\Phi}{1+\Phi}}{G \left(\beta_2 + \beta_1 \beta_4 + \frac{(1-\beta_4)\beta_1}{\Phi+1} \right) + (\beta_1 + \beta_2)(1-G)(\Phi+1)^{-\mu}} \tag{B85}$$

This expression is increasing in Φ. From (B79), Φ moves like w_H/w_L, therefore the labor share decreases in N (the opposite of w_H/w_L) when $H \approx H^P$ (this result may not extend if H^P is far from H when β_4 is close to 1).

Further, we can rewrite (B79) as:

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

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 (B85) and write:

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

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 $H \approx H^P$. ∎

B10.3 Asymptotic behavior

The asymptotic behavior is in line with Proposition 2 but the fact that automation now replaces low-skill workers with a Cobb-Douglas aggregate of capital and high-skill workers limit the ratio between the growth rate of high-skill and low-skill wages. In addition, we here need to consider the long-run behavior of the gross rental rate $\widetilde{r}$. Since r is determined by the Euler equation, then on a path where consumption growth is asymptotically constant, then $\widetilde{r}$ is also asymptotically constant (see (B74)). 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 B.6: *Consider four processes $[N_t]_{t=0}^{\infty}$, $[G_t]_{t=0}^{\infty}$, $\left[H_t^P\right]_{t=0}^{\infty}$ and $[\widetilde{r}_t]_{t=0}^{\infty}$ where $\left(N_t, G_t, H_t^P, \widetilde{r}_t\right) \in (0,\infty)\times[0,1]\times(0,H]\times(0,\infty)$ for all t. Assume that G_t, g_t^N, H_t^P and $\widetilde{r}_t$ all admit positive and finite limits with $G_\infty \in (0,1)$. Then the asymptotic growth rate of high-skill wages w_{Ht} and output Y_t are*

$$g_\infty^{w_H}=g_\infty^{Y}=g_\infty^{N}/\left[(\sigma-1)\left(\beta_2+\beta_1\beta_4\right)\right], \tag{B86}$$

and the asymptotic growth rate of low-skill wages is

$$g_\infty^{w_L}=\frac{1+(\sigma-1)\,\beta_1\beta_4}{1+(\sigma-1)\,\beta_1}g_\infty^{w_H}. \tag{B87}$$

Proof. For simplicity we assume that the limits $g_\infty^{w_H}, g_\infty^{w_L}$ and g_∞^{Y} exist (although we could show that formally as we did in Appendix A2). Suppose that $g_\infty^{w_L} \leq \beta_4 g_\infty^{w_H}$. Then Φ_t must either tend toward a positive constant or toward 0, in either case (B79) 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

(B76), we obtain:

$$w_{Ht}^{\beta_2+\beta_1\beta_4} \underset{t\to\infty}{\sim} \frac{\sigma-1}{\sigma} \frac{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}\left(G_\infty\varphi\right)^{\frac{1}{\sigma-1}}}{\left(1+\tau_m\right)^{(1-\beta_4)\beta_1}\widetilde{r}_\infty^{\beta_3+(1-\beta_4)\beta_1}} N_t^{\frac{1}{\sigma-1}}.$$

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

Moreover (B79) now implies

$$\frac{w_{Ht}H_\infty^P}{w_{Lt}L} \underset{t\to\infty}{\sim} \frac{G_\infty\left(\beta_2+\beta_1\beta_4\right)\Phi_t^\mu}{\beta_1\left(1-G_\infty\right)},$$

$$\Longrightarrow w_{Lt}^{1+\beta_1(\sigma-1)} \underset{t\to\infty}{\sim} \frac{\beta_1\left(1-G_\infty\right)\left(1+\tau_m\right)^{(1-\beta_4)\beta_1(\sigma-1)}\widetilde{r}_\infty^{(1-\beta_4)\beta_1(\sigma-1)}H_\infty^P}{G_\infty\left(\beta_2+\beta_1\beta_4\right)\varphi^\mu L} w_{Ht}^{1+\beta_4\beta_1(\sigma-1)},$$

which implies (B87). Since $\frac{1+(\sigma-1)\beta_1\beta_4}{1+(\sigma-1)\beta_1} > \beta_4$, we verify that $g_\infty^{w_L} > \beta_4 g_\infty^{w_H}$. ∎

B10.4 Dynamic equilibrium

We can solve for the dynamic equilibrium as in the baseline model. The long-run elasticity of output with respect to the number of products is now given by $\psi \equiv 1/\left[(\sigma-1)\left(\beta_2+\beta_1\beta_4\right)\right]$. We then introduce the same normalized variables as in the baseline model: $\widehat{V}_t^A$, $\widehat{V}_t^N$, $\widehat{\pi}_t^A$, $\widehat{\pi}_t^N$, $\widehat{h}_t^A$, $\widehat{c}_t$ and $\widehat{v}_t$. We also introduce $\widehat{Y}_t \equiv Y_t N^{-\psi}$ and $\widehat{K}_t \equiv K_t N^{-\psi}$. Finally we now define

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

and

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

so that

$$\left(\frac{w_L}{\widetilde{r}_t^{1-\beta_4} w_H^{\beta_4}}\right)^{\beta_1(1-\sigma)} = \omega_t n_t.$$

The transitional dynamics can then be expressed as a system of differential equations in $\mathbf{x}_t \equiv \left(n_t, G_t, \widehat{K}_t, \widehat{h}_t^A, \widehat{v}_t, \widehat{c}_t\right)$ where the first three variables are state variables and the last three control variables.

Equation (A16) still applies, therefore, we get using (B75) that

$$\pi_t^A = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\left(\varphi\left(((1+\tau_m)\widetilde{r}_t)^{1-\beta_4} w_H^{\beta_4}\right)^{1-\epsilon} + w_L^{1-\epsilon}\right)^{\mu}\left(\frac{w_H^{\beta_2}\widetilde{r}^{\beta_3}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}\right)^{1-\sigma} Y_t.$$

We can rewrite this as

(B88)
$$\widehat{\pi}_t^A = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\left(\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}\right)^{\sigma-1}\left(\widetilde{r}_t^{(1-\beta_4)\beta_1+\beta_3}\widehat{v}_t^{\beta_2+\beta_4\beta_1}\right)^{1-\sigma}\left(\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu}\widehat{Y}_t.$$

We can derive π_t^N similarly and we find

$$\widehat{\pi}_t^N = \omega_t n_t\left(\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu}\widehat{\pi}_t^A. \tag{B89}$$

(A7) is now replaced by

$$\dot{n}_t = -\frac{1-\beta_4}{1+\beta_1(\sigma-1)}\frac{\beta_1}{\beta_2+\beta_1\beta_4} g_t^N n_t. \tag{B90}$$

(A8) still applies and so does (A9). Because of the automation tax (A10) is replaced by

$$\left(r_t - (\psi-1)g_t^N\right)\widehat{V}_t^N = \widehat{\pi}_t^N + \eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa}\left(\widehat{V}_t^A - \widehat{V}_t^N\right) - (1+\tau_a)\widehat{v}_t\widehat{h}_t + \dot{\widehat{V_t^N}} \tag{B91}$$

and (A11) by

$$\kappa\eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa-1}\left(\widehat{V}_t^A - \widehat{V}_t^N\right) = (1+\tau_a)\widehat{v}_t. \tag{B92}$$

Combining (B89), (B91), (B92) and (17), we now obtain:

(B93)
$$\dot{\widehat{v}}_t = \widehat{v}_t\left(\widetilde{r}_t - \Delta - (\psi-1)g_t^N - \gamma\omega_t n_t\left(\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{-\mu}\frac{\widehat{\pi}_t^A}{\widehat{v}_t} - \gamma(1+\tau_a)\frac{1-\kappa}{\kappa}\widehat{h}_t^A\right).$$

Following the same steps as those used to derive (A14), we now obtain:

(B94)
$$\begin{aligned}\dot{\widehat{h}}_t^A &= \frac{\gamma \widehat{h}_t^A}{1-\kappa}\left(\omega_t n_t \left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{-\mu} \frac{\widehat{\pi}_t^A}{\widehat{v}_t} + \left(1+\tau_a\right)\frac{1-\kappa}{\kappa}\widehat{h}_t^A\right) \\ &\quad - \frac{\kappa\eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa}}{\left(1-\kappa\right)\left(1+\tau_a\right)}\left(1-\omega_t n_t\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{-\mu}\right)\frac{\widehat{\pi}_t^A}{\widehat{v}_t} \\ &\quad + \eta G_t^{\widetilde{\kappa}}\left(\widehat{h}_t^A\right)^{\kappa+1} + \frac{\widetilde{\kappa}}{1-\kappa}\left(\eta\left(\widehat{h}_t^A\right)^{\kappa+1} G_t^{\widetilde{\kappa}-1}\left(1-G_t\right) - g_t^N \widehat{h}_t^A\right).\end{aligned}$$

Further, (18) still applies and we can rewrite it as:

(B95)
$$\dot{\widehat{c}}_t = \frac{\widehat{c}_t}{\theta}\left(\widetilde{r}_t - \left(\rho + \Delta + \theta\psi g_t^N\right)\right).$$

Finally, we can rewrite (23) as

(B96)
$$\dot{\widehat{K}}_t = \widehat{Y}_t - \widehat{c}_t - \left(\Delta + \psi g_t^N\right)\widehat{K}_t.$$

Equations (B90), (A8), (B93), (B94), (B95) and (B96) form a system of differential equations which depend on $\widehat{Y}_t$, $\widehat{\pi}_t^A$, $\widetilde{r}_t$ and g_t^N.

(B76) implies

$$\frac{\sigma}{\sigma-1}\frac{\widehat{v}^{\beta_2+\beta_4\beta_1}\widetilde{r}^{\beta_3+\beta_1(1-\beta_4)}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}\left(G\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + \left(1-G\right)\omega_t n_t\right)^{\frac{1}{1-\sigma}} = 1,$$

so that

(B97)
$$\widetilde{r} = \left[\frac{\sigma-1}{\sigma}\frac{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}{\widehat{v}^{\beta_2+\beta_4\beta_1}}\left(G\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + \left(1-G\right)\omega_t n_t\right)^{\frac{1}{\sigma-1}}\right]^{\frac{1}{\beta_3+\beta_1(1-\beta_4)}},$$

which defines $\widetilde{r}$ as a function of $\mathbf{x}_t$ and ω_t. (B78) can be written as:

(B98)
$$H_t^P = \frac{\widetilde{r}_t\widehat{K}_t}{\widehat{v}_t}\frac{G_t\left(\beta_2 + \frac{\beta_1\beta_4\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}{(\omega_t n_t)^{\frac{1}{\mu}}+(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}\right)\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + \beta_2\left(1-G_t\right)\omega_t n_t}{G_t\left(\beta_3 + \frac{\beta_1(1-\beta_4)\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)-1}}{(\omega_t n_t)^{\frac{1}{\mu}}+\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}\right)\left(\varphi\left(1+\tau_m\right)^{(1-\beta_4)(1-\epsilon)} + \left(\omega_t n_t\right)^{\frac{1}{\mu}}\right)^{\mu} + \beta_3\left(1-G_t\right)\omega_t n_t},$$

which gives, together with (B97), H_t^P as a function of $\mathbf{x}_t$ and ω_t. g_t^N still obeys (A23), which then defines it as a function of $\mathbf{x}_t$ and ω_t.

Combine (B82), (B77), (B78) and (B83) to obtain:

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

which we can rewrite as

(B99)
$$\widehat{Y}_t = \frac{\sigma}{\sigma-1}\frac{\left[G_t\left(\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + (1-G)\,\omega_t n_t\right]\widehat{v}_t H_t^P}{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}{\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}+(\omega_t n_t)^{\frac{1}{\mu}}}\right)\left(\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + \beta_2(1-G)\,\omega_t n_t}.$$

This expression, with the previous equations, gives $\widehat{Y}_t$ as a function of $\mathbf{x}_t$ and ω_t. (B88) then ensures that $\widehat{\pi}_t^A$ is defined as a function $\mathbf{x}_t$ and ω_t.

Finally, from (B77) we obtain:

(B100)
$$\omega_t = \left[\frac{\left(\frac{\widehat{v}_t}{\widetilde{r}_t}\right)^{1-\beta_4}\frac{H_t^P}{L}\beta_1\left(G\left(\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}(\omega_t n_t)^{-\frac{1}{\mu}}+1\right)^{\mu-1} + (1-G)\right)}{G\left(\beta_2 + \frac{\beta_1\beta_4\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}{(\omega_t n_t)^{\frac{1}{\mu}}+\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)}}\right)\left(\varphi(1+\tau_m)^{(1-\beta_4)(1-\epsilon)} + (\omega_t n_t)^{\frac{1}{\mu}}\right)^{\mu} + \beta_2(1-G)\,\omega_t n_t}\right]^{\frac{\beta_1(1-\sigma)}{1+\beta_1(\sigma-1)}},$$

which implicitly defines ω_t as a function of $\mathbf{x}_t$. Hence, together with (B97), (B98), (A23), (B99), (B88) and (B100), the system formed by (B90), (A8), (B93), (B94), (B95) and (B96) describes the dynamic equilibrium. We then obtain

PROPOSITION B.7: *Assume that*

(B101)
$$\rho\left(\frac{(1+\tau_a)^{\kappa}}{\kappa^{\kappa}(1-\kappa)^{1-\kappa}\eta}\left(\frac{\rho}{\gamma}\right)^{1-\kappa} + \frac{1}{\gamma}\right) < \psi H$$

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

Proof. As before, we directly get that in a steady-state with $g^{N*} > 0$, we must have $n^* = 0$. (B100) then implies that ω^* is a constant defined by

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

This guarantees that in such a steady-state, $w_{Lt} \sim \omega^{*\frac{1}{\beta_1(1-\sigma)}} \widetilde{r}^{*1-\beta_4} \widehat{v}^{*\beta_4} N_t^{\frac{1+\beta_4\beta_1(\sigma-1)}{1+\beta_1(\sigma-1)} \frac{1}{(\sigma-1)(\beta_2+\beta_1\beta_4)}}$ such that $g_\infty^{w_L} = \frac{1+\beta_4\beta_1(\sigma-1)}{1+\beta_1(\sigma-1)} g_\infty^{w_H}$ as stipulated in Proposition B.6.

In addition, (B95) implies that in steady-state,

$$\widetilde{r}^* = \rho + \Delta + \theta \psi g^{N*}. \tag{B102}$$

(B98) implies that

$$H^{P*} = \frac{\widetilde{r}^* \widehat{K}^*}{\widehat{v}^*} \frac{\left(\beta_2 + \beta_1 \beta_4\right)}{\left(\beta_3 + \beta_1 \left(1 - \beta_4\right)\right)}. \tag{B103}$$

Then (B99) implies that

$$\widehat{Y}^* = \frac{\sigma}{\left(\sigma - 1\right) \left(\beta_2 + \beta_1 \beta_4\right)} \widehat{v}^* H^{P*}. \tag{B104}$$

We then get that (B88) implies that

$$\frac{\widehat{\pi}^{A*}}{\widehat{v}^*} = \frac{\left(\sigma - 1\right)^{\sigma-2}}{\sigma^{\sigma-1}} \frac{\left(\beta_1^{\beta_1} \beta_2^{\beta_2} \beta_3^{\beta_3}\right)^{\sigma-1}}{\left(\beta_2 + \beta_1 \beta_4\right)} \left(\widetilde{r}^{*(1-\beta_4)\beta_1+\beta_3} \widehat{v}^{*\beta_2+\beta_4\beta_1} \left(1 + \tau_m\right)^{(1-\beta_4)\beta_1} \right)^{1-\sigma} \varphi^\mu H^{P*}. \tag{B105}$$

(B97) gives

$$\widetilde{r}^* = \left[\frac{\sigma - 1}{\sigma} \frac{\beta_1^{\beta_1} \beta_2^{\beta_2} \beta_3^{\beta_3}}{\widehat{v}^{*\beta_2+\beta_4\beta_1}} \frac{\left(G^* \varphi^\mu\right)^{\frac{1}{\sigma-1}}}{\left(1 + \tau_m\right)^{(1-\beta_4)\beta_1}} \right]^{\frac{1}{\beta_3+\beta_1(1-\beta_4)}}. \tag{B106}$$

Therefore (B105) simplifies into

$$\frac{\widehat{\pi}^{A*}}{\widehat{v}^*} = \psi \frac{H^{P*}}{G^*}, \tag{B107}$$

just as in the baseline model. Then (B93) and (B102) together imply that

$$\widehat{h}^{A*} = \frac{\kappa}{\gamma \left(1 + \tau_a\right) \left(1 - \kappa\right)} \left(\rho + \left(\left(\theta - 1\right) \psi + 1\right) g^{N*}\right). \tag{B108}$$

This defines $\widehat{h}^{A*}$ as an increasing function of g^{N*}. Further, in steady-state G^* still obeys (A26) and H^{P*} obeys (A28), which imply that G^* and H^{P*} also be defined as function of g^{N*}.

(B107), (B94), (A26), (B108) then lead to

$$\text{(B109)} \qquad \frac{1-\kappa}{\kappa}\frac{\gamma G^*\left(1+\tau_a\right)\widehat{h}^{A*}}{\psi H^{P*}}\left(\frac{1+\tau_a}{\kappa\eta G_t^{\widetilde{\kappa}}}\left(\widehat{h}^{A*}\right)^{1-\kappa}+\frac{1}{\gamma}\right)=1,$$

which up to the term $1+\tau_a$ is the same as (A30) 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 (B101) is satisfied. As (B109), (A26), (A28) and (B108) are independent of τ_m, so are g^{N*}, $\widehat{h}^{A*}$ (now given by (B108)), G^* (given by (A26)) and H^{P*} (given by (A28)).

We further obtain $\widetilde{r}^*$ through (B102), which must be independent of τ_m as well. We then get $\widehat{v}^*$ through (B106) as

$$\widehat{v}^*=\left[\frac{\sigma-1}{\sigma}\frac{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}{\widetilde{r}^{*\beta_3+\beta_1(1-\beta_4)}}\frac{\left(G^*\varphi^{\mu}\right)^{\frac{1}{\sigma-1}}}{\left(1+\tau_m\right)^{(1-\beta_4)\beta_1}}\right]^{\frac{1}{\beta_2+\beta_4\beta_1}}.$$

We then get $\widehat{K}^*$ through (B103) and $\widehat{c}^*$ from (B96) which, using (B104), implies:

$$\widehat{c}^*=\frac{\sigma}{(\sigma-1)\left(\beta_2+\beta_1\beta_4\right)}\widehat{v}^*H^{P*}-\left(\Delta+\psi g^{N*}\right)\widehat{K}^*.$$

■

Further if $\tau_a=\tau_m=0$, $g^{N*},G^*,\widehat{h}^{A*}$ are determined by the same equations as in the baseline model except that the definition of ψ has changed. It is then direct that Proposition B.4 extends to this case.

B10.5 Short-run effect of a machine tax

We look at the short-run effect of a machine tax on wages, taking as given the allocation of high-skill labor between innovation and production and the total supply of capital (but not its allocation or the rental rate). Using (B80), we can rewrite (B76) and (B78) as:

$$\text{(B110)} \qquad N^{\frac{1}{1-\sigma}}\frac{\sigma}{\sigma-1}\frac{w_H^{\beta_2}\widetilde{r}^{\beta_3}w_L^{\beta_1}}{\beta_1^{\beta_1}\beta_2^{\beta_2}\beta_3^{\beta_3}}\left(G\left(\Phi+1\right)^{\mu}+1-G\right)^{\frac{1}{1-\sigma}}=1,$$

$$\text{(B111)}\qquad \frac{\widetilde{r}K}{w_H H^P} = \frac{G\left(\beta_3 + \frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^\mu + \beta_3(1-G)}{G\left(\beta_2 + \frac{\beta_1\beta_4\Phi}{1+\Phi}\right)(\Phi+1)^\mu + \beta_2(1-G)}.$$

Then, totally log-differentiate (B80), (B79) and (B111) to get:

$$\text{(B112)}\qquad \frac{1}{\varepsilon-1}\widehat{\Phi} = \widehat{w}_L - \widehat{w}_H + (1-\beta_4)\left(\widehat{w}_H - \widehat{\widetilde{r}}\right) - (1-\beta_4)\,\widehat{1+\tau_m}$$

(B113)

$$\widehat{w}_H - \widehat{w}_L = \left(\begin{array}{c} \frac{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^\mu}{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^\mu+\beta_2(1-G)}\left(\mu + \frac{\beta_1\beta_4\frac{1}{1+\Phi}}{\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}}\right) \\ +\frac{(1-\mu)G(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1}+(1-G)} \end{array}\right)\frac{\Phi\widehat{\Phi}}{\Phi+1},$$

(B114)

$$\begin{aligned}\widehat{\widetilde{r}} - \widehat{w}_H = {} & \frac{G\left(\beta_3 + \frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^\mu}{G\left(\beta_3 + \frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^\mu + \beta_3(1-G)}\left(\frac{\mu\Phi}{\Phi+1}\widehat{\Phi} + \frac{\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\left(\frac{\widehat{\Phi}}{1+\Phi} - \widehat{1+\tau_m}\right)}{\beta_3 + \frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}}\right) \\ & - \frac{G\left(\beta_2 + \frac{\beta_1\beta_4\Phi}{1+\Phi}\right)(\Phi+1)^\mu}{G\left(\beta_2 + \frac{\beta_1\beta_4\Phi}{1+\Phi}\right)(\Phi+1)^\mu + \beta_2(1-G)}\left(\mu + \frac{\beta_1\beta_4\frac{1}{1+\Phi}}{\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}}\right)\frac{\Phi\widehat{\Phi}}{1+\Phi}\end{aligned}$$

Combine (B112), (B113) and (B114) to get:

(B115)

$$\left(\begin{array}{c} \frac{1}{\varepsilon-1}\frac{1+\Phi}{\Phi} + \frac{\beta_4 G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^\mu}{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^\mu+\beta_2(1-G)}\left(\mu + \frac{\beta_1\beta_4\frac{1}{1+\Phi}}{\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}}\right) \\ +\frac{(1-\mu)G(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1}+(1-G)} + \frac{(1-\beta_4)G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^\mu}{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^\mu+\beta_3(1-G)}\left(\mu + \frac{\frac{\beta_1(1-\beta_4)}{(1+\tau_m)(1+\Phi)}}{\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}}\right) \end{array}\right)\frac{\Phi\widehat{\Phi}}{1+\Phi}$$

$$= -(1-\beta_4)\,\beta_3\frac{G(\Phi+1)^\mu + (1-G)}{G\left(\beta_3 + \frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^\mu + \beta_3(1-G)}\widehat{1+\tau_m}.$$

As $\mu < 1$, the coefficient in front of $\widehat{\Phi}$ on the LHS is positive, so that Φ is decreasing in the machine tax τ_m. Using (B113), we also get that the skill premium decreases in the machine tax.

Totally log-differentiating (B110), one gets

$$\widehat{w}_L = \frac{\mu}{\sigma - 1} \frac{G\left(\Phi+1\right)^{\mu}}{G\left(\Phi+1\right)^{\mu} + 1 - G} \frac{\Phi\widehat{\Phi}}{\Phi+1} - \left(1-\beta_1\right)\left(\widehat{w}_H - \widehat{w}_L\right) - \beta_3\left(\widehat{\widehat{r}} - \widehat{w}_H\right).$$

Plugging (B113) and (B114) and (B115), we get:

$$\begin{aligned} &\widehat{w}_L \\ &= -\left(\begin{array}{c} \frac{(1-\mu)G(\Phi+1)^{\mu-1}}{G(\Phi+1)^{\mu-1}+(1-G)}\left[1-\beta_1+\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+(1-G)}\frac{\beta_1\Phi}{(1+\tau_m)(1+\Phi)}\right]-\frac{\beta_1}{\varepsilon-1}\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+1-G}\frac{\Phi\widehat{\Phi}}{\Phi+1}\frac{\tau_m}{1+\tau_m} \\ +\frac{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}}{G\left(\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}\right)(\Phi+1)^{\mu}+\beta_2(1-G)}\left(\mu+\frac{\beta_1\beta_4\frac{1}{1+\Phi}}{\beta_2+\beta_1\beta_4\frac{\Phi}{1+\Phi}}\right)\left[\beta_2+\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+(1-G)}\frac{\beta_4\beta_1\Phi}{(1+\tau_m)(1+\Phi)}\right] \\ +\frac{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}}{G\left(\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}\right)(\Phi+1)^{\mu}+\beta_3(1-G)}\left(\mu+\frac{\frac{\beta_1(1-\beta_4)}{(1+\tau_m)(1+\Phi)}}{\beta_3+\frac{\beta_1(1-\beta_4)\Phi}{(1+\tau_m)(1+\Phi)}}\right)\left[\beta_3+\frac{G(\Phi+1)^{\mu}}{G(\Phi+1)^{\mu}+(1-G)}\frac{(1-\beta_4)\beta_1\Phi}{(1+\tau_m)(1+\Phi)}\right] \end{array} \right)\frac{\Phi\widehat{\Phi}}{\Phi+1}. \end{aligned}$$

For a small tax on machines (i.e. around $\tau_m = 0$), the only negative term inside the parenthesis drops out, so that the introduction of a small tax on machines leads to an increase in low-skill wages. Therefore we have established:

PROPOSITION B.8: *On impact, a tax on machines reduces the skill premium and a small tax on machines increases low-skill wages.*

B11. Appendix to the Quantitative Exercise

Section B11.1 describes the calibration technique and Section B11.2 the data used for the calibration and the patent data presented in Figure 1.C and 7. Section B11.3 discusses how parameters are identified. Section B11.4 presents details on the constant G alternative model discussed in Section III.B. Section B11.5 shows that the data still require an increase in the share parameter G when there is constant labor-augmenting technical change. Finally, Section B11.6 carries an analysis of the effect of automation taxes in our calibrated model.

B11.1 Technique

We choose parameters to minimize the log-deviation of predicted and observed variables for the four time paths of the skill-premium, the labor-share of GDP, stock of equipment over GDP and an index of GDP per hours worked. That is, for a given set of parameters b the model produces predicted output of $\hat{\mathbf{Y}}_i = \left\{\hat{\mathbf{Y}}_{i,t}\right\}_{t=1}^{T_i}$ for each of these four paths from 1963 and until 2012 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 $\hat{\mathbf{Y}}(b) = \{\hat{\mathbf{Y}}_i(b)\}_{i=1}^{4}$ as the combined vector of these paths and make explicit the dependency on the parameters b. $\mathbf{Y}$ is the corresponding vector of actual values. We then solve:

$$min_b(log(\hat{\mathbf{Y}}(b)) - log(\mathbf{Y}))'\mathbf{W}(log(\hat{\mathbf{Y}}(b)) - log(\mathbf{Y})),$$

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 in the previous working paper and the one pursued here.

B11.2 Data

Calibration Data. We do not seek to match the skill-ratio H/L but take it as exogenously given. We normalize $H + L = 1$, throughout. The skill-ratio is taken from Acemoglu and Autor (2011). However, since our estimation requires a skill-ratio both before and after the period 1963-2007 we match the observed path of the log of the skill-ratio to a "generalized" logistical function of the form:

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

where (α, β, μ, s) are parameters to be estimated. We use the observed skill-ratio in the period $1963 - 2007$ and the predicted values outside of this time interval. Yet, the fit is so good that there is no visual difference in the match of the four time periods between this approach and using the predicted value in the interval $1963 - 2007$. The skill-premium is taken from Autor (2014) which extends the data of the Acemoglu and Autor (2011) until 2012. The labor share is the BLS's labor share in the non-farm business sector (BLSb, 2020). We take GDP per hours worked from the series on non-farm business from the BLS (BLSa, 2020). Capital equipment is calculated as follows: We follow KORV and use quality-adjusted price indices of equipment from Cummins and Violante (2002) who update the series from Gordon (1990). We combine two different series. First, we use NIPA data on private investment in equipment excluding transportation equipment (Tables 1.5 and 5.3.5. from NIPA, see BEA, 2020). We iteratively construct an index for the stock of private real capital equipment by assuming a depreciation rate of 12.5 per cent (as Krusell et al., 2000) and using the price index for private equipment from Cummins and Violante (2002). We start this approach in 1947 but only use the stock from 1963 onwards. We combine this with the growth rate of real private GDP to get an index for equipment over GDP.

We match this index to the NIPA private equipment capital stock (excluding transportation) over (private) GDP number for 1963 to get a series in absolute value. To this, we add software, but following the suggestion of Cummins and Violante (2002), we use the NIPA data on the stock of software over GDP (table 2.1 from NIPA). We add these two values to get our combined stock of equipment (+software) over GDP.

Labor costs data. We combine three sources of data. First, two indices from FRED (Federal Reserve Bank of St. Louis, 2020), the *Average Hourly Earnings of Production and Nonsupervisory Employees, Total Private (AHETPI_NBD20120101)* which contains the wages and salaries of workers and Nonfarm Business Sector and *Compensation Per Hour (COMPNFB_NBD20120101)*, which contains total compensation of workers. These indices give the trend growth. To couple this with a level, we use the BLS Bureau of Labor Statistics from March 1991 which gives wages and salaries as a share of total compensation of 72.3 per cent for private workers in the United States for 1991. We combine these to gives a time path since 1964 of share of total compensation going to wages and salaries (we equate 1963 to 1964). This path is consistent with other more recent time trends for the BLS. For instance, the December 2019 release of the Employment Cost Index from the BLS shows a small decline in the share of total compensation going to wages of .7 per cent from 2005 to 2019 (combining tables 4 and 8) which is consistent with our measure which shows a small increase of 1 per cent over the same time period. Finally, though the share of total compensation not going to wages and salaries is not constant across occupational groups it is substantial for all groups and is never below 27 per cent for full-time workers (https://www.bls.gov/news.release/ecec.t05.htm).

Patent data on automation innovations. The data for Figure 7.A are taken from Mann and Püttmann (2018) who classify USPTO patents granted from 1976. Given that they classify patents according to their grant years, we lag all their numbers by 2 years to reproduce an approximate time lag of 2 years between application year (which is closer to the year of innovation) and grant year. They find that commuting zones exposed to industries with a higher level of automation experience a decline in manufacturing employment (and an increase in overall employment).

Dechezleprêtre et al. (2019) classify patents in machinery as automation versus non-automation. Their classification method follows two steps. First, they classify technological codes (IPC and CPC codes, mostly at the 6 digit level) by computing the frequency of certain keywords which have been related to automation (such as "robot", "automation", "computer numerical control", etc.) for each technological code in machinery. They identify automation technological codes as those with a high share of patents with a keyword (in the top 5 percent of the distribution) and non-automation codes as those with a low share (in the bottom 60 percent). Second, they define automation patent as patents with at least one automation technological code and non-automation patents as patents with only non-automation codes. They show that the share of automation patents in machinery in a sector is correlated with a decline in routine manual and cognitive tasks, and an increase in the high-skill to low-skill employment ratio. Then, they use cross-country variation in wages and variations in firms' exposures to different countries, to show in firm-level regressions, that an increase in low-skill wages leads to an increase in automation innovations but not in non-automation innovations. They classify patents using the PATSTAT database which starts before 1976. In Figure 7.B, we use granted patents at the USPTO.

B11.3 Parameters identification

In this section, we discuss how our parameters are identified, first by carrying a back-of-the-envelope calibration, second by computing the elasticities of the initial and final values of the series we match with respect to the parameters, and third by computing how precisely each parameter is identified. We then discuss specifically how $\tilde{\kappa}$ is determined and finally carry an out-of-sample prediction exercise, where we only use the first 30 years of the data to calibrate our parameters.

Back-of-the-envelope calibration. We first study how the production parameters σ, β_1, β_2, β_4 and Δ would be identified under a naive back-of-the-envelope calibration, where we assume that in 1963 the U.S. economy was in the first phase while in 2012, it was in the third phase. Since both assumptions are actually not met in our estimation, this naive calibration gives parameters that are still far from those which we actually estimate. Nevertheless, the exercise is informative to understand how these production parameters are related to moments in the data.

Assuming that the economy in 1963 is close to the first phase, and using (B77), we get that the skill premium must obey:

$$\frac{w_{H1963}}{w_{L1963}} \approx \frac{\beta_2}{\beta_1} \frac{L_{1963}}{H_{1963} - \frac{1}{\gamma} g_{1963}^N}.$$

Further, using that most high-skill workers work in production, such that $\frac{1}{\gamma} g_{1963}^N$ is small relative to H, we obtain

$$\frac{\beta_2}{\beta_1} \approx \frac{w_{H1963} H_{1963}}{w_{L1963} L_{1963}}, \tag{B116}$$

so that the ratio β_2/β_1 is determined by the ratio between the high-skill wage bill and the low-skill wage bill. Because the economy is in fact not in the first phase in 1963 (with an equipment stock to GDP ratio which is not 0), this approximation is likely to overstate the ratio β_2/β_1. Similarly, using (B77), (B78), and (B84), we get that the labor share in 1963 should obey

$$ls_{1963} \approx \frac{\beta_2 \frac{H}{H-\frac{1}{\gamma}g^N_{1963}} + \beta_1}{\frac{\sigma}{\sigma-1} + \beta_2 \frac{\frac{1}{\gamma}g^N_{1963}}{H-\frac{1}{\gamma}g^N_{1963}}},$$

which simplifies into

$$ls_{1963} \approx \frac{\sigma-1}{\sigma}\left(\beta_2+\beta_1\right), \tag{B117}$$

if most high-skill workers are in production. Therefore, given σ, the initial labor share determines β_3, the 'external' capital share. We can then combine (B116) and (B117) to obtain

$$\beta_1 \approx \frac{1}{\frac{w_{H1963}H_{1963}}{w_{L1963}L_{1963}}+1}\frac{ls_{1963}}{1-\frac{1}{\sigma}}, \tag{B118}$$

so that β_1 which is the Cobb-Douglas share for low-skill workers in Phase 1 is given by the labor share in 1963 and the ratio between the high-skill wage bill and the low-skill wage bill, and σ which determines mark-ups.

Combining (B87) and (B87), we get that if the economy is close to its asymptotic steady-state in 2012, the growth rate of the skill premium is given by

$$g^{sp}_{2012} \approx \frac{\beta_1\left(\sigma-1\right)\left(1-\beta_4\right)}{1+\beta_1\left(\sigma-1\right)} g^{GDP}_{2012}. \tag{B119}$$

Using (B77), (B78) (B84), the labor share now obeys:

$$ls_{2012} \approx H\left[\frac{\sigma H}{\left(\sigma-1\right)\left(\beta_2+\beta_1\beta_4\right)} - \left(\frac{\sigma}{\left(\sigma-1\right)\left(\beta_2+\beta_1\beta_4\right)}-1\right)\left(\frac{1}{\gamma}g^N_{1963}+H^A_{1963}\right)\right]^{-1},$$

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

$$ls_{2012} \approx \frac{\sigma-1}{\sigma}\left(\beta_2+\beta_1\beta_4\right). \tag{B120}$$

Combining (B116), (B117) and (B120) we obtain:

$$\beta_4 \approx 1-\left(1-\frac{ls_{2012}}{ls_{1963}}\right)\left(\frac{w_{H1963}H_{1963}}{w_{L1963}L_{1963}}+1\right). \tag{B121}$$

Therefore, in this approximation, β_4 is identified through the decline in the labor share and the initial wage bill ratio between high-skill and low-skill workers. In the data the labor share does not monotonically decline. To understand how the parameters are identified, we replace ls_{2012} by the lowest value over 1963-2012 (which is 57%) and ls_{1963} by the highest value (64.6%). With $\frac{w_{H1963}H_{1963}}{w_{L1963}L_{1963}} = 0.576$, we then obtain $\beta_4 \approx 0.82$. This is higher than the value we actually end up finding ($\beta_4 = 0.73$), mostly because the economy is still far from its steady-state in 2012 (so that ls_{2012} is higher than the asymptotic value of the labor share).

Using (B118), (B119) and (B121) we obtain:

$$\sigma \approx \frac{1}{ls_{1963}\left[\frac{g_{2012}^{GDP}}{g_{2012}^{sp}}\left(1-\frac{ls_{2012}}{ls_{1963}}\right)-\frac{1}{\frac{w_{H1963}H_{1963}}{w_{L1963}L_{1963}}+1}\right]}.$$

that is given the initial wage bill ratio and the labor shares in 1963 and 2012, which inform us about β_1, β_2 and β_4, σ is determined by the ratio between the growth rate of GDP and that of the skill-premium in the third phase. The larger is σ, the more automated firms gain over non-automated ones and therefore the more the skill premium rises relative to GDP: hence a lower $\frac{g_{2012}^{GDP}}{g_{2012}^{sp}}$ is associated with a larger σ. When using the last 10 years to determine $\frac{g_{2012}^{GDP}}{g_{2012}^{sp}}$, we find that $\sigma \approx 6.77$, while our estimation procedure leads to $\sigma = 5.94$.

Given σ, one can then find β_1 using (B118), we find $\beta_1 \approx 0.48$, below but not too far from the estimated value of 0.59 (this approximation is not too sensitive on σ provided that σ is large enough). Using (B116), we then obtain $\beta_2 \approx 0.28$ which is higher than the estimated value of 0.18, in line with the fact that (B116) gives an overestimate of β_2/β_1.

To get a proxy for Δ, we look at the steady-state value for the equipment to GDP ratio. Using (B103), (B104) and the definition of GDP, we obtain that

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

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

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

since in steady-state the economy is Cobb-Douglas with a total physical capital share of $\beta_3+\beta_1(1-\beta_4)$ and an equipment of share $\beta_1(1-\beta_4)$. Using (B102), we obtain

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

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

$$\rho + \Delta + \theta g_{2000}^{GDP} = \left(\frac{\sigma - 1}{\sigma}\right) \frac{\beta_1 (1-\beta_4)}{\frac{K_{eq,2000}}{GDP_{2000}}} \approx 0.051 \tag{B122}$$

using the values computed above. It is therefore not surprising that we find a low Δ in the estimation. This is due in particular to the high-level of $\frac{K_{eq,2000}}{GDP_{2000}} = 1.5$ (with the actual estimated values for σ, β_1 and β_4 we would still find that $\rho + \Delta + \theta g_{2000}^{GDP} \approx 0.088$).

Parameters' effect on empirical moments. We now show the role that parameters have on the empirical moments which allows us to identify what features of the data pin down the parameters. Taking as our starting point the parameter estimates of Section III, we iteratively change each one by 2% and show the resulting effects on the initial (1963) and the final (2012) values of each of the four empirical paths. Table B3 reports the elasticities (note that β_3 is completely determined by β_1 and β_2).

The initial skill premium is most strongly affected by the production function parameters β_1, β_2 and β_4: A higher share of high-skill workers in production, β_2, directly increases the skill-premium. A higher value of β_4 makes automation more expensive, which increases the demand for low-skill workers and reduces the skill premium. A higher β_1 implies a lower β_3 which reduces the role of structural capital. This reduces the rental rate of capital, which increases the use of capital equipment and thereby the skill-premium. β_2 has the opposite effect on the skill premium in 2012. A higher β_2 reduces the multiplier of N_t on output, Y_t which reduces the growth rate of the economy. The automation technology parameters, κ, $\tilde{\kappa}$, η also have a large effect on the skill premium in 2012.

The initial labor share depends on β_1, β_2 and β_4, the latter having a much larger effect in 2012 since the share of automated products is much larger.

$GDP/labor$ is mechanically affected negatively by higher σ since we keep the stock of products in 1963 constant. Both β_1 and β_2 reduce the importance of structural capital and thereby have a negative effect on $GDP/labor$ in 1963 as the stock of capital is sufficiently large. In 2012 $\sigma, \beta_1, \beta_2, \beta_4$ all reduce the multiplier of N_t on Y_t and therefore $GDP/labor$. The innovation parameters γ, η lead to higher growth and therefore higher $GDP/labor$ in 2012, though naturally not in 1963.

Capital equipment / GDP in 1963 depends positively on β_1 and negatively on β_4 because the initial capital stock is fixed. For 2012, a higher β_4 increases the cost of automation and thereby reduces K_{eq}/GDP. Horizontal innovation productivity, γ, encourages more innovation. This drives up the wage of high-skill workers in 1963, makes automation more expensive and reduces K_{eq}/GDP. It further increases the growth rate of the economy and reduces G_{2012} such that K_{eq}/GDP in 2012 is also lower. Finally, higher productivity of machines, $\tilde{\varphi}$, shifts capital into equipment and consequently raises K_{eq}/GDP.

Table B3—The effect of parameters on the four empirical paths (numbers refer to elasticities of empirical value wrt. parameter)

Parameters	Skill premium 1963	Skill premium 2012	Labor share 1963	Labor share 2012	GDP/labor 1963	GDP/labor 2012	K_{eq}/GDP 1963	K_{eq}/GDP 2012
σ	−0.1	0.0	0.1	0.0	−0.5	−1.9	2.1	1.1
ϵ	−0.1	−0.1	0.0	0.0	0.0	−0.2	−0.3	−0.2
β_1	0.3	−0.1	0.7	0.6	−1.0	−2.6	7.7	1.8
γ	0.1	−0.1	0.0	0.1	0.0	0.7	−0.2	−0.5
$\tilde{\kappa}$	−0.2	−0.2	0.0	0.1	0.0	0.0	0.3	−0.4
θ	0.0	−0.1	0.0	0.0	0.0	−0.3	0.1	−0.2
η	0.1	0.3	0.0	0.0	0.0	0.2	−0.1	0.2
κ	−0.1	−0.2	0.0	0.0	0.0	−0.3	0.3	−0.2
ρ	−0.1	−0.2	0.0	0.0	0.0	−0.6	0.2	−0.3
$\tilde{\varphi}$	0.1	0.3	0.0	0.0	0.0	0.1	0.5	0.2
β_2	0.4	−0.2	0.2	0.3	−0.6	−1.5	1.1	−0.1
Δ	0.0	−0.1	0.0	0.0	0.0	−0.1	0.0	−0.2
β_4	−0.7	−1.9	0.1	0.6	−0.1	−1.3	−5.0	−3.7
N_{1963}	0.0	0.1	0.0	0.0	0.2	0.3	−0.2	0.0
G_{1963}	0.1	0.0	0.0	0.0	0.0	−0.1	0.6	0.0

Precision of the parameters. In the following, we calculate the effect the parameters have on the aggregate final moment. We do this allowing for all the other parameters to adjust, illustrating how precisely each of the parameters are determined. Since deviations from the minimum parameter values are naturally second order we do not compute elasticities. Instead, for a given parameter θ_i consider

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

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

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

We compute the expression on the left. The results are in Table B4 for a 2% shock on the parameter of interest. It shows that the parameters that govern the production function: $(\sigma,\beta_1,\beta_2,\beta_4)$ are the hardest to vary and consequently the ones most precisely identified. The exception is ϵ, the elasticity between low-skill labor and machines, which as Proposition 2 makes clear, does not govern the asymptotic growth of income inequality. ρ,θ,η,γ all govern the growth rate of the economy and are weakly identified individually. Increases in φ can to a certain extend be accommodated by changes to N_{1963} and consequently neither is very well-identified. The depreciation of capital Δ is also not well identified because it mostly affects the growth rate of the capital stock which also depends on ρ and θ (equation B122). Given that this parameter is the one estimated outside a common range this is a reassuring finding.

Table B4—The "curvature" of deviating from the optimal parameter

Parameters	σ	ϵ	β_1	γ	$\tilde{\kappa}$	θ	η	κ
Curvature (x 10^4	16.0	0.5	44.0	0.7	0.4	0.8	0.6	0.9
Parameters	ρ	$\tilde{\varphi}$	β_2	Δ	β_4	N_{1963}	G_{1963}	
Curvature	1.0	0.8	2.4	0.5	428.1	0.5	0.4	

B11.4 Details on the constant G calibration

Table B5 shows the calibrated parameters in the alternative model described in section III.B where G is an exogenous constant. To match the data as well as possible, this constant G model requires very high elasticity of substitution across intermediates σ and between low-skill workers and machines in automated firms ϵ. The share of automated products is estimated at $G = 0.9$. Figure B14 reports the two moments not mentioned in section III.B. This model captures well the trend in labor productivity, but performs worse for the equipment stock to GDP ratio than the baseline model.

Table B5—Parameters from quantitative exercise for a constant G model

Parameter	σ	ϵ	β_1	γ	θ	ρ
Value	10.6	39.1	0.51	1.03	1	0.012
Parameter	β_2	Δ	β_4	$\tilde{\varphi}$	N_{1963}	G
Value	0.17	0.007	0.84	1.51	0.16	0.90

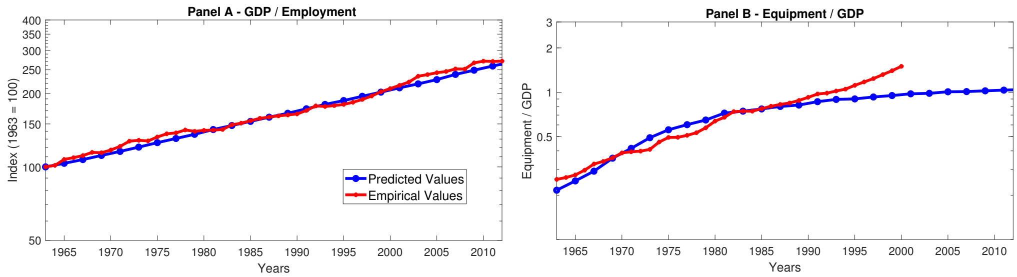

Figure B14. Predicted and empirical time paths for a model with constant G.

B11.5 Labor-augmenting technical change

In Section III.B, we showed that conditional on our production function, the data require a path for G_t similar to that generated by our endogenous growth model. To assess how robust that result is, we add high-skill labor augmenting technical change to our model. That is, we replace (22) with:

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

where A_{Ht} is high-skill labor augmenting technical change with g_{A_H} a constant. We then look for the exogenous paths for N_t and G_t together with an initial value A_{H1963} and a growth rate g_{A_H} which best match the data (still assuming the baseline parameters of Table 1). We find that $A_{H1963} = 1.025$ and a very small $g_{A_H} = 0.03\%$. The resulting path for G_t is still very similar to the one generated by the endogenous growth model as illustrated by Figure B15. Note that allowing instead for exogenous low-skill labor augmenting technical change or machine augmenting technical change (an exogenous trend on $\tilde{\varphi}$) is isomorphic to this exercise for the optimal G_t path (each exercise would deliver a different optimal N_t path).

B11.6 Automation taxes

Among the many policy proposals to address rising income inequality, is a tax on the use of automation technology or a "robot tax". Here, we analyze two distinct taxes: on the use of machines—in the form of a tax on the rental rate of equipment— and on the innovation of new machines—in the form of taxing high-skill workers in automation innovation—see Appendix B10 for details. In either cases we consider the permanent unexpected introduction of a 25% tax in the first non-calibrated year, 2013.

First, consider a tax on the use of machines. To clarify the role of endogenous technology we also simulate the economy holding technology, N_t and G_t and therefore H_t^P at the baseline level. Figure B16 reports the results. The immediate effect is to discourage the use of machines and consequently low-skill wages rise

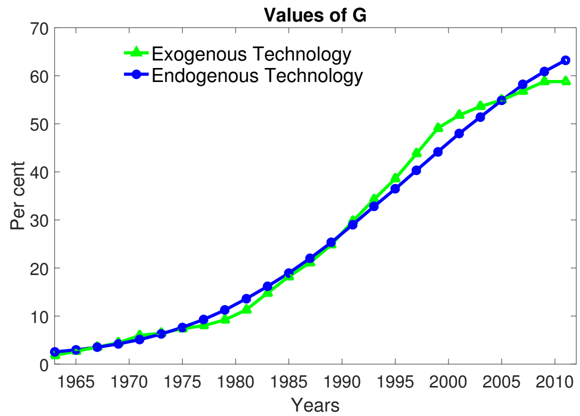

FIGURE B15. PATH G_tIN THE ENDOGENOUS GROWTH MODEL AND IN THE MODIFIED EXOGENOUS GROWTH MODEL WITH HIGH-SKILL AUGMENTING TECHNICAL CHANGE.

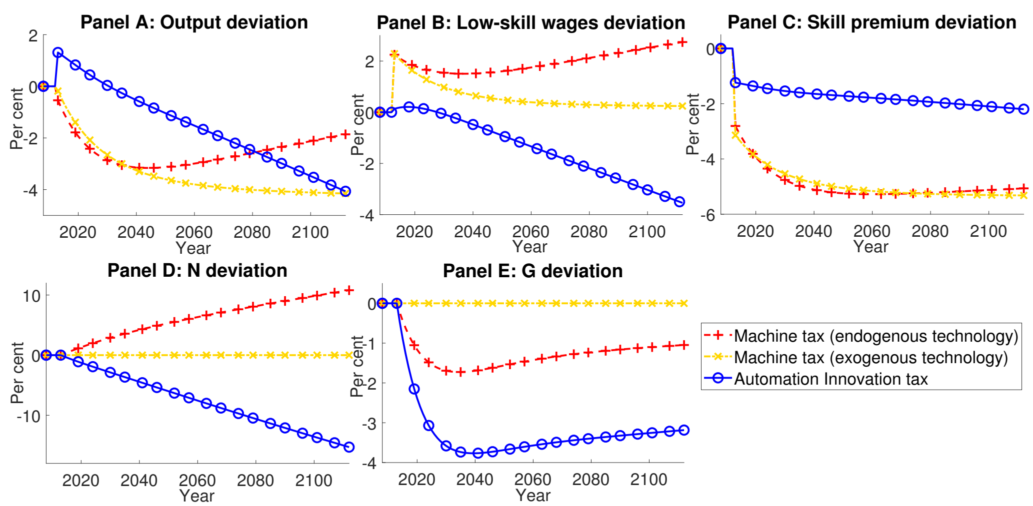

FIGURE B16. EFFECTS OF A MACHINE TAX AND AN AUTOMATION INNOVATION TAX RELATIVE TO BASELINE by 2% on impact (Panel B) with a corresponding lower skill premium (Panel C) (In Appendix B10.5 we show that low-skill wages will increase on impact for any parameter values.[44]) The endogeneity of technology amplifies the effect of the tax over time (in panel B, the gap

between the endogenous and the exogenous cases widens). This results from two effects. First, the tax discourages automation innovation leading to a lower G (Panel E). Second, since high-skill workers and machines are complements, the tax reallocates high-skill workers away from production and toward horizontal innovation, increasing N (Panel D). Consequently, the positive effect on low-skill wages is eventually larger than the initial 2%. Output initially decreases on impact in a similar fashion whether technology is endogenous or not, but it recovers and eventually (beyond the horizon of the figure) increases in the endogenous case (Panel A) due to the increase in N_t.[45]

A tax on automation innovation has very different implications: First, high-skill workers move from innovation in automation to production which, on impact, boosts output and marginally low-skill wages. As the share of automated products G_t decreases, low-skill wages further modestly increase. However, discouraging automation innovation also discourages horizontal innovation since the not-yet automated firms are the ones bearing the burden of the tax. This eventually reduces low-skill wages. The intuition is similar to that of Proposition B.4 since a tax on automation innovation has similar effects to reducing the effectiveness of the automation technology. Quantitatively, the effect remains modest since it takes 30 years for the number of products to decrease by 5% (which correspond to a decrease of 0.17 p.p. in annual growth rate). The skill-premium is also reduced as the economy grows at a slower rate.

This exercise highlights the importance of endogenous technology: Though both forms of "robot" taxes increase low-skill wages on impact, the long-run effects depend crucially on whether the tax is designed to encourage or discourage overall innovation. Of course, this exercise is only a first pass and analyzing the welfare consequences of these policies or others, say minimum wage legislation, is of interest for future research.

B12. Comparison with KORV

We show formally the claims made in Section III.A that KORV cannot replicate a decline in the labor share without other counterfactual predictions and does not feature labor-saving innovation. Using their notation, their production function

[44] By comparison in KORV, the effect of such a tax depends on parameters.

[45] Asymptotically, a machine tax has no effect on G or on the growth rate of N: as using low-skill workers instead of machines becomes prohibitively expensive, the allocation of high-skill workers remains undistorted by the presence of a finite tax. As a result, in the long-run, G_t reaches the same steady-state but N_t is at a permanently higher level because for a long time the tax has created excess horizontal innovation. See Proposition B.7 in Appendix B10.

is given by:

$$F = Ak_s^\alpha \left(\mu u^\sigma + (1-\mu)\left(\lambda k_e^\rho + (1-\lambda)\, s^\rho\right)^{\sigma/\rho}\right)^{\frac{1-\alpha}{\sigma}}, \tag{B123}$$

where k_s is structure, u is low-skill labor, s is high-skill labor and k_e is equipment. The key features are that $\sigma > \rho$ and k_e increases faster than GDP.

In their estimation, k_e and h are strict complements ($\rho < 0$), so as k_e keeps increasing because of investment-specific technological change, its factor share must eventually go to 0; meaning that the long-run prediction of their model is an *increase* in the labor share. Even though, their estimation rejects $\rho \geq 0$, it is worth checking what happens in that case. If equipment and high-skill workers are substitutes ($\rho > 0$, which is the calibrated parameter in Eden and Gaggl, 2018), the economy experiences explosive growth (which seems counterfactual) since in the long-run it becomes an AK model where K/Y grows from investment specific technical change. If $\rho = 0$, then their production function looks like the one of our automated products, and indeed the capital share must eventually increase. But, then, the growth rate of the skill premium is given by:[46]

$$g_{\pi t} = (1-\sigma)(g_u - g_s) + \sigma\lambda(g_{k_{et}} - g_s).$$

Consequently, if investment specific technological change accelerates (that is there are relatively more and more innovations of that type such that $g_{k_{et}}$ grows), then the skill premium must grow faster (this is also the case for $\sigma > \rho > 0$). This parameterization will now have problems with the first puzzle that we solve: namely a slow down in the growth rate in the skill premium at a time where technical change is the most directed toward "automation" / investment specific technical change.[47]

We now show that investment-specific technical change is not low-skill labor saving in KORV. To do so, we solve for the low-skill wage in their model and consider an increase in investment specific technical change q_t. q_t is the extra TFP parameter in the production of equipment investment compared to the consumption

[46]Here we use their equation (4) $g_{\pi t} = (1-\sigma)(g_u - g_s) + (\sigma - \rho)\lambda\left(\frac{k_{et}}{s_t}\right)^\rho (g_{k_{et}} - g_s)$, leaving out the labor augmenting terms, which are not included in their preferred specification and do not reflect capital-skill complementarity. This does not affect the present point.

[47]KORV briefly discuss a production function where the nests are inverted so that:

$$F = Ak_s^\alpha \left(\mu s^\sigma + (1-\mu)\left(\lambda k_e^\rho + (1-\lambda)\, u^\rho\right)^{\sigma/\rho}\right)^{\frac{1-\alpha}{\sigma}} \tag{B124}$$

with $\sigma < \rho$. This specification does not match their data but is similar to our specification within automated firm (but not for the aggregate economy). Here again the same issues arise: if $\sigma < 0$, then the long-run capital share declines. If $\sigma > 0$, growth is explosive. If $\sigma = 0$ and $\rho > 0$, the capital share increases in the long-run but the skill premium cannot grow less fast when technical change is the most directed toward investment.

good (so that $1/q_t$ is the price of the investment good for equipment). We look here at the effect of a one time permanent increase in q_t, keeping the expected price change $E_t\left(\frac{q_t}{q_{t+1}}\right)$ fixed and assuming a fixed rental rate on structures (or equivalently a fixed interest rate) $R_{st} = r_t + \delta_s$. Note that we need to make such assumptions because KORV do not specify a supply function for capital (since the capital stock is simply taken from the data). This assumption corresponds to a perfectly elastic capital stock which is how we evaluate the one time effect of a change in G_t in Proposition 1, in the discussion of the effect of automation on wages in Section III and in Proposition B.5 in Appendix B10. Taking first order condition in (B123), we get the rental rate on structures:

$$R_{st} = \alpha k_{st}^{\alpha-1}\left(\mu u_t^{\sigma} + (1-\mu)\left(\lambda k_{et}^{\rho} + (1-\lambda)\, s_t^{\rho}\right)^{\frac{\sigma}{\rho}}\right)^{\frac{1-\alpha}{\sigma}}. \tag{B125}$$

KORV assume that the returns on both capital stocks must be equal, that is:

$$1 - \delta_s + R_{st} = E_t\left(\frac{q_t}{q_{t+1}}\right)(1-\delta_e) + q_t R_{et}, \tag{B126}$$

which implies that the rental rate on equipment obeys:

$$R_{et} = \frac{1}{q_t}\left(1 - \delta_s + R_{st} - E_t\left(\frac{q_t}{q_{t+1}}\right)(1-\delta_e)\right).$$

Therefore R_{et} decreases with q_{et}. Taking first order condition in (B123) with respect to k_{et}, and using (B125), we get:

$$R_{et} = (1-\mu)\,\lambda\left(\lambda + (1-\lambda)\frac{s_t^{\rho}}{k_{et}^{\rho}}\right)^{\frac{1-\rho}{\rho}}(1-\alpha)\left(\frac{\alpha}{R_{st}}\right)^{\frac{\alpha}{1-\alpha}}\left(\mu\frac{u_t^{\sigma}}{\left(\lambda k_{et}^{\rho} + (1-\lambda)\, s_t^{\rho}\right)^{\frac{\sigma}{\rho}}} + (1-\mu)\right)^{\frac{1-\sigma}{\sigma}},$$

which shows that (as expected) k_{et} decreases in R_{et} so that k_{et} increases if q_t increases. Finally, the first order condition with respect to unskilled labor is given by

$$w_{Lt} = (1-\alpha)\, k_{st}^{\alpha}\mu u_t^{\sigma-1}\left(\mu u_t^{\sigma} + (1-\mu)\left(\lambda k_{et}^{\rho} + (1-\lambda)\, s_t^{\rho}\right)^{\frac{\sigma}{\rho}}\right)^{\frac{1-\alpha}{\sigma}-1}. \tag{B127}$$

Combining this with (B125) gives

$$w_{Lt} = (1-\alpha)\left(\frac{\alpha}{R_{st}}\right)^{\frac{\alpha}{1-\alpha}}\mu u_t^{\sigma-1}\left(\mu u_t^{\sigma} + (1-\mu)\left(\lambda k_{et}^{\rho} + (1-\lambda)\, s_t^{\rho}\right)^{\frac{\sigma}{\rho}}\right)^{\frac{1-\sigma}{\sigma}}. \tag{B128}$$

Therefore an increase in q_t leads to an increase in k_{et} and consequently low-skill wages w_{Lt}: investment specific technical change is not labor saving in KORV's main specification.[48] This is true regardless of the parameters σ and ρ (and therefore also applies to Eden and Gaggl, 2018).

ADDITIONAL REFERENCE

Nordhaus, William. 2007. "Two Centuries of Productivity Growth in Computing." *Journal of Economic History*, 67(1): 128 - 159.

Federal Reserve Bank of St. Louis and US. Office of Management and Budget. 2020. "*Average Hourly Earnings of Production and Nonsupervisory Employees"* and *"Compensation Per Hour". https://fred-redesign.stlouisfed.org*

Bureau of Labor Statistics, 2020a, U.S. Bureau of Labor Statistics, Nonfarm Business Sector: Labor Productivity (Output per Hour) for All Employed Persons [PRS85006092], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PRS85006092, October 9, 2020

Bureau of Labor Statistics. 2020b. "Labor's share of output in the nonfarm business sector" downloaded from https://www.bls.gov/opub/mlr/2017/images/data/giandrea-sprague-fig1.stm

Bureau of Economic Analysis. 2020. "Fixed Assets Accounts Tables". Table 1.5 from https://apps.bea.gov/iTable/iTable.cfm?ReqID=10&step=2 plus other tables listed.

[48]It can be labor-saving with the alternative nesting given in (B124).