Online Appendix for the paper
"The Race Between Man and Machine:
Implications of Technology for Growth, Factor Shares and Employment,"
by Daron Acemoglu and Pascual Restrepo.

Appendix B (Not-For-Publication): Omitted Proofs and Additional Results

Details of the Empirical Analysis

This section provides information about the data used in constructing Figures 1 and 9. We also provide a regression analysis documenting the robustness of the patterns illustrated in these figures.

Data: We use data on employment counts for 304 occupational categories that we can track consistently over time, from 1980 until 2015. Our occupational categories roughly match the 330 categories proposed by David Dorn (see `http://www.ddorn.net/data.htm`). We aggregate some of these categories to account for merged occupational codes in recent waves of the American Community Survey. The details of our approach can be found in the replication files that accompany this paper.

We use data from the Census for 1980, 1990, and 2000, as well as the American Community Survey for 2010 and 2015 (Ruggles et al. 2017). Using these data, we compute for each of our 304 occupational categories the total employment count and the demographic characteristics of its workers, including their gender, age, education, race and whether they are foreign born (we focus on workers between 16 and 64 years of age). We also compute the share of jobs in each occupational category that are in manufacturing, the primary sector (agriculture, forestry, fisheries and mining), and services (retail trade, finance, business and repair services, personal services, entertainment services, professional services, and public administration).

Our measure of new job titles comes from Lin (2011), who computes the total amount of job titles and new job titles in each occupational category for 1980, 1990 and 2000.[34] Lin identifies new job titles by comparing changes across waves of the *Dictionary of Occupational Titles*, and also by comparing the *1990 Census Index of Occupations* with its 2000 counterpart. Importantly, Lin uses official documentation to avoid labeling as new those jobs that were simply reclassified or divided because of reasons unrelated to the type of work people performed (i.e., because of administrative changes in U.S. statistical agencies). Instead, Lin's measure counts a job as new if workers perform a different set of tasks in this job than in any previously existing jobs. The data on new and total job titles can be matched consistently to 303 of our occupations in 1980 and 1990, and to all of our occupations in 2000.

Detailed Analysis for Figure 1: In addition to Figure 1 in the main text, in Figure B1 we pool the 1980-2015 changes together with the 1990-2015 and 2000-2015 changes. In this case, the

[34]The data are available from Jeffrey Lin's website `https://sites.google.com/site/jeffrlin/newwork`

share of new job titles refers to this variable measured at the beginning of each time window (i.e., 1980, 1990 or 2000). A very similar positive relationship is visible in the figure.

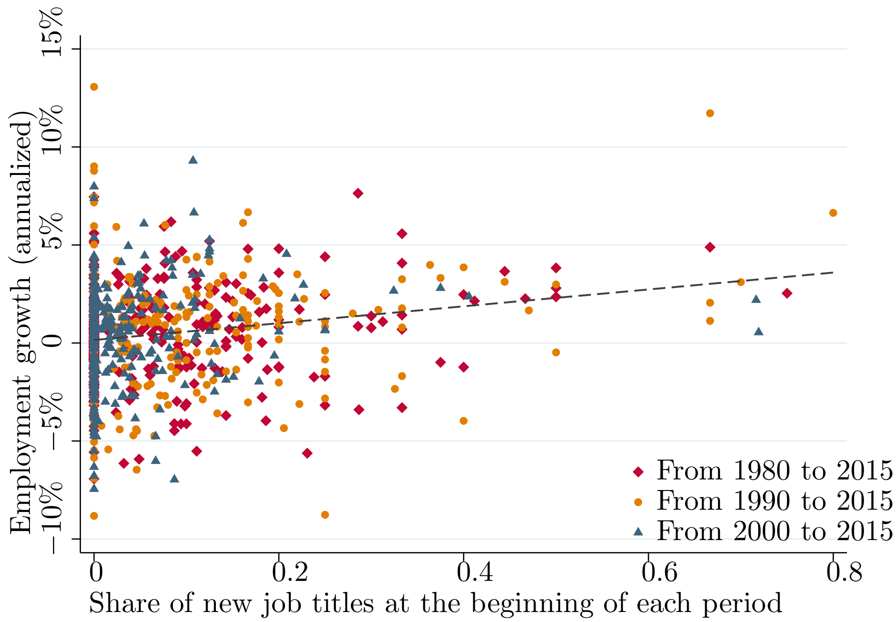

FIGURE B1: Employment growth by occupation over different time periods (annualized), plotted against the share of new job titles at the beginning of each period in each occupation.

We further document this relationship and probe its robustness by estimating the following regression:

$$\Delta \ln E_{it} = \beta N_{it} + \delta_t + \Phi_t X_{it} + u_{it}. \tag{B1}$$

Here, the dependent variable is the (annualized) growth in employment in occupation i. The key explanatory variable is N_{it}—the share of new job titles in occupational category i at the beginning of the period.

We start in Table B1 with the 1980-2015 change as in Figure 1. In this case, there is only one observation per occupation, and we report standard errors that are robust against heteroscedasticity.

Column 1 shows the raw correlation without any covariates, which is positive and statistically significant.

Column 2 includes the initial level of employment and total number of job titles in each occupation. This leads to a larger and more precisely estimated coefficient on the share of new job titles: 3.953 (standard error = 1.080). Column 3, which is our baseline specification shown in Figure 1, in addition controls for the demographic composition of employment in each occupation—in particular, allowing for differential growth by average age, fraction male, share foreign-born, fraction black and fraction Hispanic in the occupation in 1980. Now the coefficient on the share of new job titles is 4.153 (standard error = 1.143). Using this estimate, we compute that if there had been no additional employment growth in occupations with more new job titles in 1980, total employment

Table B1: Long-differences estimates of employment growth in occupational categories with a higher baseline share of new job titles.

	Dependent variable: percent change in employment growth between 1980-2015 (annualized).					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of new job titles in 1980	2.602**	3.953***	4.153***	3.425***	3.254***	3.334***
	(1.264)	(1.080)	(1.143)	(1.059)	(1.014)	(1.146)
log of employment in 1980		-0.001	-0.002	-0.002	-0.003**	-0.003***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log of job titles in 1980		-0.705***	-0.647***	-0.484***	-0.272*	-0.128
		(0.112)	(0.120)	(0.114)	(0.156)	(0.121)
Average age of workers in 1980			-0.055	-0.063	-0.022	-0.042
			(0.044)	(0.044)	(0.042)	(0.038)
Share of male workers in 1980			-1.035*	-0.484	0.270	0.151
			(0.540)	(0.533)	(0.601)	(0.547)
Share of foreign workers in 1980			2.725	-8.628*	1.455	2.887
			(4.838)	(4.685)	(4.786)	(4.759)
Share of black workers in 1980			-2.904	1.531	-2.044	-2.350
			(2.931)	(2.635)	(2.367)	(2.706)
Share of Hispanic workers in 1980			-3.964	17.642***	2.959	1.887
			(5.913)	(5.888)	(5.873)	(5.872)
Average years of education in 1980				0.562***		
				(0.101)		
Share of manufacturing jobs in 1980					-1.756**	-2.342***
					(0.698)	(0.696)
Share of primary sector jobs in 1980					-0.012	-0.393
					(0.699)	(0.904)
Share of service sector jobs in 1980					2.160***	1.831***
					(0.575)	(0.584)
R-squared	0.01	0.18	0.20	0.27	0.32	0.31
Observations	303	303	303	303	303	303
Covariates and estimation:						
Total job titles and occupation size		✓	✓	✓	✓	✓
Demographics			✓	✓	✓	✓
Educational level				✓		
Sectoral composition					✓	✓
Robust regression						✓

Notes: The table presents long-difference estimates of the relationship between the share of new job titles in an occupational category in 1980 and subsequent employment growth between 1980-2015 (annualized). The table also reports the coefficients estimated for the covariates included in each model. Finally, in column 6 we present robust-regression estimates following Li (1985). Standard errors that are robust against heteroscedasticity are presented in parentheses.

growth between 1980 and 2015 would have been 24% rather than 60%. This is the basis of our claim in the text that about 60% of employment growth between 1980 and 2015 is associated with faster employment growth in occupations with more new job titles.

In column 3, we do not control for average education in the occupation, since, as we discuss further below, occupations with more new job titles attract more educated workers, making this variable a "bad control."Nevertheless, column 4 shows that controlling for it does not alter the qualitative relationship between share of new job titles and employment growth, though the coefficient now declines modestly to 3.425 (standard error = 1.059).

In column 5 we add share of manufacturing, primary sector and service job titles in each occupation in 1980 to the specification of column 3. These variables enable us to control for the general structural change in the economy away from manufacturing and primary sector jobs towards service jobs. This also leads to a somewhat lower estimate, which still remains precisely estimated: 3.254 (standard error = 1.014). Finally, in column 6, we estimate a robust regression down-weighing outliers and excessively influential observations (following Li, 1985). The results are very similar.

In Table B2, we estimate the same models now exploiting variation in the share of new job titles at the beginning of each decade between 1980 and 2000. Panel A looks at a sample consisting of stacked differences for 1980-2015, 1990-2015 and 2000-2015. Panel B is for 1980-2010, 1990-2010 and 2000-2010. Finally, Panel C focuses on decadal changes, 1980-1990, 1990-2000 and 2000-2010. In each case, the share of new job titles refers to this variable measured at the beginning of the period for the relevant time window. In addition, we control for a full set of period dummies and the standard errors are now robust against heteroscedasticity and serial correlation at the occupation level. The results are very similar to those reported in Table B1 in all cases.

Table B2: Stacked-differences estimates of employment growth in occupational categories with a higher baseline share of new job titles.

	Dependent variable: percent change in employment growth (annualized).					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: stacked differences for 1980-2015, 1990-2015, and 2000-2015.					
Share of new job titles at t	4.284***	4.399***	4.448***	3.499***	3.562***	3.150***
	(0.894)	(0.881)	(0.909)	(0.954)	(1.021)	(0.696)
R-squared	0.05	0.16	0.18	0.23	0.29	0.30
Observations	910	910	910	910	910	910
	Panel B: stacked differences for 1980-2010, 1990-2010, and 2000-2010.					
Share of new job titles at t	4.297***	4.490***	4.593***	3.428***	3.537***	3.310***
	(1.009)	(0.968)	(1.003)	(1.047)	(1.122)	(0.712)
R-squared	0.04	0.17	0.20	0.25	0.34	0.36
Observations	910	910	910	910	910	910
	Panel C: stacked differences for 1980-1990, 1990-2000, and 2000-2010.					
Share of new job titles at t	3.809**	4.223**	4.277**	3.088*	3.224*	3.406***
	(1.804)	(1.683)	(1.737)	(1.849)	(1.839)	(0.866)
R-squared	0.04	0.13	0.16	0.19	0.23	0.35
Observations	910	910	910	910	910	910
Covariates and estimation:						
Decadal dummies	✓	✓	✓	✓	✓	✓
Total job titles and occupation size		✓	✓	✓	✓	✓
Demographics			✓	✓	✓	✓
Educational level				✓		
Sectoral composition					✓	✓
Robust regression						✓

Notes: The table presents stacked-differences estimates of the relationship between the share of new job titles in an occupational category and subsequent employment growth (annualized). In Panel A, we stack the data for the periods from 1980-2015, 1990-2015 and 2000-2015. In Panel B, we stack the data for the periods from 1980-2010, 1990-2010 and 2000-2010. In Panel C, we stack the data for the three decades 1980-1990, 1990-2000 and 2000-2010. All models include a full set of decadal effects. In addition, we introduce the covariates indicated at the bottom rows, but do not report their coefficients. In column 2 we control for the log of total job titles and employment in each occupational category. In column 3 we also control for the demographic characteristics of workers in each occupation (average age, gender, place of birth, and race). In column 4 we further add average years of schooling in the occupation. In column 5 we control for the sectoral composition of jobs in each occupation (share of jobs in manufacturing, the primary sector, and services). All these covariates are allowed to have time-varying coefficients. Finally, in column 6 we present robust-regression estimates following Li (1985). Standard errors that are robust against heteroscedasticity and serial correlation within occupations are presented in parentheses.

Detailed Analysis for Figure 9: We now briefly present regression evidence documenting that the pattern shown in Figure 9 is robust. In particular, we report estimates of the following equation

$$H_{it} = \beta N_{it} + \delta_t + \Phi_t X_{it} + u_{it}, \tag{B2}$$

where the left-and side variable is the average years of schooling (or the share of workers with college) among workers employed in occupation i at time t, while N_{it} is again the share of new job titles in occupational category i at time t. The regressions always include period dummies and standard errors are robust against heteroscedasticity and serial correlation at the occupation level. The five columns in Table B3 correspond to columns 1-3 and 5-6 of Table B1 (because the left-hand side variable is average years of schooling, we do not control for it on the right-hand side). The results show that in all specifications there is a significant positive association between the share of new job titles in an occupation and the average years of schooling of workers in the subsequent decades. The relationship shown in Figure 9 corresponds to column 3, where we control for differential demographic trends.

Table B3: Estimates of the education level of workers in occupational categories with a higher baseline share of new job titles.

	Dependent variable: educational attainment of workers within each occupational category.				
	(1)	(2)	(3)	(4)	(5)
	Panel A: Average years of education among workers.				
Share of new job titles at t	3.144***	3.082***	1.685***	1.324***	1.221***
	(0.522)	(0.543)	(0.363)	(0.335)	(0.304)
R-squared	0.05	0.17	0.67	0.71	0.75
Observations	910	910	910	910	910
	Panel B: Share of workers with a college degree.				
Share of new job titles at t	0.420***	0.414***	0.215***	0.160**	0.131**
	(0.085)	(0.088)	(0.068)	(0.065)	(0.053)
R-squared	0.03	0.12	0.55	0.60	0.64
Observations	910	910	910	910	910
Covariates and estimation:					
Decadal dummies	✓	✓	✓	✓	✓
Total job titles and occupation size		✓	✓	✓	✓
Demographics			✓	✓	✓
Sectoral composition				✓	✓
Robust regression					✓

Notes: The table presents estimates of the relationship between the share of new job titles in an occupational category and the educational level of its workers. In these estimates, we pool data from 1980, 1990 and 2000. In Panel A, we use the average years of schooling among workers in an occupational category as the dependent variable. In Panel B, we use the share of workers with college degree as dependent variable. All models include a full set of decadal effects. In addition, we introduce the covariates indicated at the bottom rows, but do not report their coefficients. In column 2 we control for the log of total job titles and employment in each occupational category. In column 3 we also control for the demographic characteristics of workers in each occupation (average age, gender, place of birth, and race). In column 4 we also control for the sectoral composition of jobs in each occupation (share of jobs in manufacturing, the primary sector, and services). All these covariates are allowed to have time-varying coefficients. Finally, in column 5 we present robust-regression estimates following Li (1985). Standard errors that are robust against heteroscedasticity and serial correlation within occupations are presented in parentheses

Remaining Proofs from Section 2

We start with the proof of Lemma A1.

Proof of Lemma A1. The assumption that $K < \overline{K}$ and $I^* \leq \widetilde{I}$ implies that:

$$\frac{W}{\gamma(N)} < R \leq \frac{W}{\gamma(I^*)}. \tag{B3}$$

We first show that $\omega(I^*, N, K)$ is (strictly) decreasing in I^*. To do so, we compute $\omega_I^*(I^*, N, K)$ and show that Assumption $2'$ is sufficient to ensure it is negative.

Log-differentiating equations (A3) and (A4), we have

$$\varepsilon_K \frac{d \ln R}{dI^*} = \frac{d \ln Y}{dI^*} + \frac{1}{I^* - N + 1}, \tag{B4}$$

$$\varepsilon_L \frac{d \ln W}{dI^*} = \frac{d \ln Y}{dI^*} - \xi(I^*), \tag{B5}$$

where

$$\begin{aligned} \varepsilon_K &= \zeta + (\sigma - \zeta)\varsigma_K, \\ \varepsilon_L &= \int_{I^*}^{N} \xi(i)(\zeta + (\sigma - \zeta)\varsigma_L(i))di, \end{aligned}$$

and $\varsigma_K \in [0, 1]$ is the share of capital in tasks produced with capital, $\varsigma_L(i) \in [0, 1]$ is the share of labor in task i, and $\xi(i) \in [0, 1]$ is the share of total payments to labor earned by workers in task i (in particular, we have $\int_{I^*}^{N} \xi(i)di = 1$).

Differentiating equation (A5), we get

$$\frac{1}{1-\sigma} \frac{c^u(W/\gamma(I^*))^{1-\sigma} - c^u(R)^{1-\sigma}}{B^{1-\sigma}} = s_K \frac{d \ln R}{dI^*} + s_L \frac{d \ln W}{dI^*}. \tag{B6}$$

where $s_K = \frac{RK}{Y} \in [0, 1]$ and $s_L = \frac{WL}{Y} \in [0, 1]$ are respectively the capital and the labor shares in national income.

Solving the system of equations determined by (B4), (B5), and (B6) yields

$$\begin{aligned} \frac{\omega_{I^*}(I^*, N, K)}{\omega(I^*, N, K)} &= \frac{d \ln W}{dI^*} - \frac{d \ln R}{dI^*}, \\ &= -\frac{s_L + s_K}{\varepsilon_K s_L + \varepsilon_L s_K} \left(\frac{1}{I^* - N + 1} + \xi(I^*) \right) \\ &\quad + \frac{\varepsilon_K - \varepsilon_L}{\varepsilon_K s_L + \varepsilon_L s_K} \frac{1}{1-\sigma} \frac{c^u(W/\gamma(I^*))^{1-\sigma} - c^u(R)^{1-\sigma}}{B^{1-\sigma}}. \end{aligned}$$

Therefore, $\omega(I^*, N, K)$ is (strictly) decreasing in I^* if and only if

$$(\varepsilon_K - \varepsilon_L) \frac{1}{1-\sigma} \frac{c^u(W/\gamma(I^*))^{1-\sigma} - c^u(R)^{1-\sigma}}{B^{1-\sigma}} < (s_L + s_k) \left(\frac{1}{I^* - N + 1} + \xi(I^*) \right).$$

Let $\varsigma_{max} = \max_{i \in [I^*, N]}\{\varsigma_L(i)\}$ and $\varsigma_{min} = \min_{i \in [I^*, N]}\{\varsigma_L(i)\}$. Inequality (B3) implies that $\varsigma_K \in [\varsigma_{min}, \varsigma_{max}]$. Thus:

$$\varepsilon_K - \varepsilon_L = (\sigma - \zeta) \left(\varsigma_K - \int_{I^*}^{N} \xi(i)\varsigma_L(i)di \right) < |\sigma - \zeta|(\varsigma_{max} - \varsigma_{min}).$$

In addition, $s_L + s_K > \varsigma_{min}$, because the share of capital or labor in every task is at least ς_{min}. Thus, the inequality

$$|\sigma - \zeta| \frac{\varsigma_{max} - \varsigma_{min}}{\varsigma_{min}} \frac{1}{1-\sigma} \frac{c^u(W/\gamma(I^*))^{1-\sigma} - c^u(R)^{1-\sigma}}{B^{1-\sigma}} < \frac{1}{I^* - N + 1} + \xi(I^*) \tag{B7}$$

suffices to ensure that $\omega(I^*, N, K)$ is (strictly) decreasing in I^*.

We now show that Assumption $2'$ implies (B7).

If $\eta \to 0$, then $\varsigma_{max} = \varsigma_{min} = 1$ and (B7) holds. Likewise, if $\zeta = 1$, $\varsigma_{max} = \varsigma_{min} = 1 - \eta$ and (B7) holds. To complete the proof we show that (B7) holds under (A1). This follows from the following sequence of inequalities:

- If $\zeta < 1$,

$$\begin{aligned}
\frac{\varsigma_{max} - \varsigma_{min}}{\varsigma_{min}} &= \frac{\frac{(W/\gamma(I^*))^{1-\zeta}}{\eta\psi^{1-\zeta} + (1-\eta)(W/\gamma(I^*))^{1-\zeta}} - \frac{(W/\gamma(N))^{1-\zeta}}{\eta\psi^{1-\zeta} + (1-\eta)(W/\gamma(N))^{1-\zeta}}}{\frac{(W/\gamma(N))^{1-\zeta}}{\eta\psi^{1-\zeta} + (1-\eta)(W/\gamma(N))^{1-\zeta}}}, \\
&= \frac{((W/\gamma(I^*))^{1-\zeta} - (W/\gamma(N))^{1-\zeta})\eta\psi^{1-\zeta}}{(W/\gamma(N))^{1-\zeta}(\eta\psi^{1-\zeta} + (1-\eta)(W/\gamma(I^*))^{1-\zeta})}, \\
&< \frac{(W/\gamma(I^*))^{1-\zeta} - (W/\gamma(N))^{1-\zeta}}{(W/\gamma(N))^{1-\zeta}}, \\
&= \left(\frac{\gamma(N)}{\gamma(I^*)}\right)^{1-\zeta} - 1, \\
&< \left(\frac{\gamma(N)}{\gamma(N-1)}\right)^{1-\zeta} - 1.
\end{aligned}$$

If, on the other hand, if $\zeta > 1$,

$$\begin{aligned}
\frac{\varsigma_{max} - \varsigma_{min}}{\varsigma_{min}} &= \frac{\frac{(W/\gamma(N))^{1-\zeta}}{\eta\psi^{1-\zeta} + (1-\eta)(W/\gamma(N))^{1-\zeta}} - \frac{(W/\gamma(I^*))^{1-\zeta}}{\eta\psi^{1-\zeta} + (1-\eta)(W/\gamma(I^*))^{1-\zeta}}}{\frac{(W/\gamma(I^*))^{1-\zeta}}{\eta\psi^{1-\zeta} + (1-\eta)(W/\gamma(I^*))^{1-\zeta}}}, \\
&= \frac{((W/\gamma(N))^{1-\zeta} - (W/\gamma(I^*))^{1-\zeta})\eta\psi^{1-\zeta}}{(W/\gamma(I^*))^{1-\zeta}(\eta\psi^{1-\zeta} + (1-\eta)(W/\gamma(N))^{1-\zeta})}, \\
&< \frac{(W/\gamma(N))^{1-\zeta} - (W/\gamma(I^*))^{1-\zeta}}{(W/\gamma(I^*))^{1-\zeta}}, \\
&= \left(\frac{\gamma(I^*)}{\gamma(N)}\right)^{1-\zeta} - 1, \\
&< \left(\frac{\gamma(N-1)}{\gamma(N)}\right)^{1-\zeta} - 1.
\end{aligned}$$

Thus,

$$\frac{\varsigma_{max} - \varsigma_{min}}{\varsigma_{min}} < \left(\frac{\gamma(N)}{\gamma(N-1)}\right)^{|1-\zeta|} - 1. \tag{B8}$$

- The function $f(x) = \frac{1}{1-\sigma}x^{1-\sigma}$ is concave. Because $\frac{W}{\gamma(I^*)} \geq R$, we also have

$$\begin{aligned}\frac{1}{1-\sigma}\frac{c^u(W/\gamma(I^*))^{1-\sigma} - c^u(R)^{1-\sigma}}{B^{1-\sigma}} &< \frac{c^u(R)^{-\sigma}(c^u(W/\gamma(I^*)) - c^u(R))}{B^{1-\sigma}},\\ &< \frac{c^u(R)^{1-\sigma}}{B^{1-\sigma}}\frac{c^u(W/\gamma(I^*))}{c^u(R)},\\ &< \frac{c^u(R)^{1-\sigma}}{B^{1-\sigma}}\frac{\gamma(N)}{\gamma(N-1)}.\end{aligned}$$

  In the last line we used the fact that:

$$\frac{c^u(W/\gamma(I^*))}{c^u(R)} < \frac{c^u(W/\gamma(N-1))}{c^u(W/\gamma(N))} \leq \frac{\gamma(N)}{\gamma(N-1)},$$

  which follows from observing that $c^u(x)/x$ is decreasing in x.

  Finally, the ideal price index condition in equation (A5) implies that

$$\frac{c^u(R)^{1-\sigma}}{B^{1-\sigma}} < \frac{c^u(R)^{1-\sigma}}{(I^* - N + 1)c^u(R)^{1-\sigma}} < \frac{1}{I^* - N + 1} + \xi(I^*).$$

  This inequality implies

$$\frac{1}{1-\sigma}\frac{c^u(W/\gamma(I^*))^{1-\sigma} - c^u(R)^{1-\sigma}}{B^{1-\sigma}} < \left(\frac{1}{I^* - N + 1} + \xi(I^*)\right)\frac{\gamma(N)}{\gamma(N-1)}. \tag{B9}$$

- Multiplying inequalities (A1), (B8), and (B9), we obtain the sufficient condition (B7). This shows that Assumption 2$'$ implies (B7), and ensures that $\omega_I(I^*, N, K)$ is negative.

We now show that $\omega(I^*, N, K)$ is (strictly) increasing in N. To do so, we compute $\omega_N(I^*, N, K)$ and show that Assumption 2$'$ is sufficient to ensure it is positive.

Log-differentiating equations (A3), (A4), and (A5), and solving for the change in wages and rental rates, we have

$$\begin{aligned}\frac{\omega_N(I^*, N, K)}{\omega(I^*, N, K)} =& \frac{d\ln W}{dN} - \frac{d\ln R}{dN},\\ =& \frac{s_L + s_K}{\varepsilon_K\lambda_L + \varepsilon_L\lambda_K}\left(\frac{1}{I^* - N + 1} + \xi(N)\right)\\ &+ \frac{\varepsilon_K - \varepsilon_L}{\varepsilon_K s_L + \varepsilon_L s_K}\frac{1}{1-\sigma}\frac{c^u(R)^{1-\sigma} - c^u(W/\gamma(N))^{1-\sigma}}{B^{1-\sigma}}.\end{aligned}$$

Therefore, $\omega(I^*, N, K)$ is (strictly) increasing in N if and only if

$$(\varepsilon_L - \varepsilon_K)\frac{1}{1-\sigma}\frac{c^u(R)^{1-\sigma} - c^u(W/\gamma(N))^{1-\sigma}}{B^{1-\sigma}} < (s_L + s_K)\left(\frac{1}{I^* - N + 1} + \xi(N)\right).$$

Inequality (B3) implies that $\varsigma_K \in [\varsigma_{min}, \varsigma_{max}]$. Thus

$$\varepsilon_L - \varepsilon_K = (\sigma - \zeta)\left(\int_{I^*}^{N} \xi(i)\varsigma_L(i)di - \varsigma_K\right) < |\sigma - \zeta|(\varsigma_{max} - \varsigma_{min}).$$

In addition, $s_L + s_K > \varsigma_{min}$, because the share of capital or labor in every task is at least ς_{min}. Thus,the inequality

$$|\sigma - \zeta|\frac{\varsigma_{max} - \varsigma_{min}}{\varsigma_{min}} \frac{1}{1-\sigma} \frac{c^u(R)^{1-\sigma} - c^u(W/\gamma(N))^{1-\sigma}}{B^{1-\sigma}} < \left(\frac{1}{I^* - N + 1} + \xi(N)\right) \tag{B11}$$

suffices to ensure that $\omega(I^*, N, K)$ is (strictly) increasing in N.

We now show that Assumption $2'$ implies (B11).

If $\eta \to 0$ then $\varsigma_{max} = \varsigma_{min} = 1$ and (B11) holds. Likewise, if $\zeta = 1$, $\varsigma_{max} = \varsigma_{min} = 1 - \eta$ and (B11) holds. To complete the proof we show that (B11) holds under (A1). This follows from the next three steps:

- Following the same steps as before, we have that (B8) holds.
- The function $f(x) = \frac{1}{1-\sigma} x^{1-\sigma}$ is concave. Because $\frac{W}{\gamma(N)} < R$,

$$\begin{aligned}\frac{1}{1-\sigma} \frac{c^u(R)^{1-\sigma} - c^u(W/\gamma(N))^{1-\sigma}}{B^{1-\sigma}} &< \frac{c^u(W/\gamma(N))^{-\sigma}(c^u(R) - c^u(W/\gamma(N)))}{B^{1-\sigma}}, \\ &< \frac{c^u(R)^{1-\sigma}}{B^{1-\sigma}} \frac{c^u(R)^{\sigma}}{c^u(W/\gamma(N))^{\sigma}}, \\ &< \frac{c^u(R)^{1-\sigma}}{B^{1-\sigma}} \left(\frac{\gamma(N)}{\gamma(N-1)}\right)^{\sigma},\end{aligned}$$

  where the last inequality follows because $c^u(x)/x$ is decreasing and thus

$$\frac{c^u(R)}{c^u(W/\gamma(N))} < \frac{c^u(W/\gamma(N-1))}{c^u(W/\gamma(N))} \leq \frac{\gamma(N)}{\gamma(N-1)}.$$

- Finally, the ideal price index condition in equation (A5) implies

$$\frac{c^u(R)^{1-\sigma}}{B^{1-\sigma}} < \frac{c^u(R)^{1-\sigma}}{(I^* - N + 1)c^u(R)^{1-\sigma}} < \frac{1}{I^* - N + 1} + \xi(N).$$

  This inequality implies

$$\frac{1}{1-\sigma} \frac{c^u(W/\gamma(I^*))^{1-\sigma} - c^u(R)^{1-\sigma}}{B^{1-\sigma}} < \left(\frac{1}{I^* - N + 1} + \xi(N)\right) \left(\frac{\gamma(N)}{\gamma(N-1)}\right)^{\sigma}. \tag{B12}$$

- Multiplying inequalities (A1), (B8), and (B12), we obtain the sufficient condition (B11). This shows that Assumption $2'$ implies (B11), and ensures that $\omega_N(I^*, N, K)$ is positive.

■

Proof of Proposition 2: We first formulate a more general version of this proposition, which holds under Assumption $2'$, and then derive the tighter characterization presented in the text (under Assumption 2). In this proof, $\frac{\partial \omega}{\partial I^*}$, $\frac{\partial \omega}{\partial N}$ and $\frac{\partial \omega}{\partial K}$ denote the partial derivatives of the function $\omega(I^*, N, K)$ with respect to its arguments.

Proposition B1 (Comparative statics in the general model) *Suppose that Assumptions 1, 2′ and 3 hold. Let $\varepsilon_L > 0$ denote the elasticity of the labor supply schedule $L^s(\omega)$ with respect to ω; let $\varepsilon_\gamma = \frac{d\ln\gamma(I)}{dI} > 0$ denote the semi-elasticity of the comparative advantage schedule.*

- *If $I^* = I < \widetilde{I}$—so that the allocation of tasks to factors is constrained by technology—then:*
  - *the impact of technological change on relative factor prices is given by*

$$\frac{d\ln(W/R)}{dI} = \frac{d\ln\omega}{dI} = \frac{1}{\omega}\frac{\partial\omega}{\partial I^*} < 0,$$
$$\frac{d\ln(W/R)}{dN} = \frac{d\ln\omega}{dN} = \frac{1}{\omega}\frac{\partial\omega}{\partial N} > 0$$

  - *the impact of capital on relative factor prices is given by*

$$\frac{d\ln(W/R)}{d\ln K} = \frac{d\ln\omega}{d\ln K} + 1 = \frac{1+\varepsilon_L}{\sigma_{cons}+\varepsilon_L} > 0,$$

    where $\sigma_{cons} \in (0,\infty)$ is the elasticity of substitution between labor and capital that applies when technology constraints the allocation of factors to tasks. This elasticity is given by a weighted average of σ and ζ.

- *If $I^* = \widetilde{I} < I$—so that the allocation of tasks to factors is cost-minimizing—then*
  - *the impact of technological change on relative factor prices is given by*

$$\frac{d\ln(W/R)}{dI} = \frac{d\ln\omega}{dI} = 0,$$
$$\frac{d\ln(W/R)}{dN} = \frac{d\ln\omega}{dN} = \frac{\sigma_{cons}+\varepsilon_L}{\sigma_{free}+\varepsilon_L}\frac{1}{\omega}\frac{\partial\omega}{\partial N} > 0,$$

  - *and the impact of capital on relative factor prices is given by*

$$\frac{d\ln(W/R)}{d\ln K} = \frac{d\ln\omega}{d\ln K} + 1 = \left(\frac{1+\varepsilon_L}{\sigma_{free}+\varepsilon_L}\right) > 0,$$

    where

$$\sigma_{free} = (\sigma_{cons}+\varepsilon_L)\left(1 - \frac{1}{\omega}\frac{\partial\omega}{\partial I^*}\frac{1}{\varepsilon_\gamma}\right) - \varepsilon_L > \hat{\sigma};$$

- *In both parts of the proposition, the labor share and employment move in the same direction as ω.*

- *Finally, under Assumption 2, we have*

$$\frac{1}{\omega}\frac{\partial\omega}{\partial I^*} = -\frac{1}{\hat{\sigma}+\varepsilon_L}\Lambda_I \qquad\qquad \frac{1}{\omega}\frac{\partial\omega}{\partial N} = \frac{1}{\hat{\sigma}+\varepsilon_L}\Lambda_N,$$

  and the elasticities of substitution are

$$\sigma_{cons} = \hat{\sigma} \qquad\qquad \sigma_{free} = \hat{\sigma} + \frac{1}{\varepsilon_\gamma}\Lambda_I.$$

Note: In this proposition, we do not explicitly treat the case in which $I^* = I = \widetilde{I}$ in order to save on space and notation, since in this case left and right derivatives with respect to I are different.

Proof. We first establish the comparative statics of ω with respect to I, N and K when both $I^* = I < \widetilde{I}$ and $I^* = \widetilde{I} < I$.

Comparative statics for K: The curve $I^* = \min\{I, \widetilde{I}\}$ does not depend on K, all comparative statics are determined by the effect of capital on $\omega(I^*, N, K)$. An increase in K shifts up the relative demand locus in Figure A1 (this does not affect the ideal price index condition, which simplifies the analysis in this case), and thus increases W and reduces R . The impact on $\omega = \frac{W}{RK}$ depends on whether the initial effect on W/R has elasticity greater than one (since K is in the denominator).

Notice that the function $\omega(I^*, N, K)$ *already* incorporates the equilibrium labor supply response. To distinguish this supply response from the elasticity of substitution determined by factor demands, we define $\omega^L(I^*, N, K, L)$ as the static equilibrium for a fixed level of the labor supply L.

The definition of σ_{cons} implies that $\frac{\partial \omega^L}{\partial K}\frac{K}{\omega^L} = \frac{1}{\sigma_{\text{cons}}} - 1$ and $-\frac{\partial \omega^L}{\partial L}\frac{L}{\omega^L} = \frac{1}{\sigma_{\text{cons}}}$. Thus, when $I^* = I < \widetilde{I}$, we have

$$d\ln(W/R) = d\ln\omega + 1 = \left(\frac{1}{\sigma_{\text{cons}}} - 1\right) d\ln K - \frac{1}{\sigma_{\text{cons}}}\varepsilon_L d\ln\omega + d\ln K = \frac{1+\varepsilon_L}{\sigma_{\text{cons}} + \varepsilon_L} d\ln K,$$

where we have used the fact that $\omega(I^*, N, K) = \omega^L(I^*, N, K, L^s(\omega))$. This establishes the claims about the comparative statics with respect to K when $I^* = I < \widetilde{I}$.

For the case where $I^* = \widetilde{I} < I$, we have that the change in K also changes the threshold task $I^* = \widetilde{I}$. In particular, $dI^* = \frac{1}{\varepsilon_\gamma} d\ln\omega$. Thus,

$$d\ln(W/R) = \frac{1+\varepsilon_L}{\sigma_{cons} + \varepsilon_L} d\ln K + \frac{1}{\omega}\frac{\partial\omega}{\partial I^*}\frac{1}{\varepsilon_\gamma} d\ln(W/R) = \frac{1+\varepsilon_L}{\sigma_{cons} + \varepsilon_L}\frac{1}{1 - \frac{1}{\omega}\frac{\partial\omega}{\partial I}\frac{1}{\varepsilon_\gamma}} d\ln K = \frac{1+\varepsilon_L}{\sigma_{\text{free}} + \varepsilon_L} d\ln K,$$

where we define σ_{free} as in the proposition.

Comparative statics with respect to I**:** The relative demand locus $\omega = \omega(I^*, N, K)$ does not directly depend on I . Thus, the comparative statics are entirely determined by the effect of changes in I on the $I^* = \min\{I, \widetilde{I}\}$ schedule depicted in Figure 3. When $I^* = \widetilde{I} < I$, small changes in I have no effect as claimed in the proposition. Suppose next that $I^* = I < \widetilde{I}$. In this case, an increase in I shifts the curve $I^* = \min\{I, \widetilde{I}\}$ to the right in Figure 3. Lemma A1 implies that $\omega(I^*, N, K)$ is decreasing in I^*. Thus, the shift in I increases I^* and reduces ω—as stated in the proposition. Moreover, because $I^* = I$, we have

$$\frac{d\ln(W/R)}{dI} = \frac{d\ln\omega}{dI^*} = \frac{1}{\omega}\frac{\partial\omega}{\partial I^*} < 0,$$

where $\frac{\partial\omega}{\partial I^*}$ denotes the partial derivative of $\omega(I^*, N, K)$ with respect to I^*.

Comparative statics for N: From Lemma A1, changes in N only shift the relative demand curve up in Figure 3. Hence, when $I^* = I < \widetilde{I}$, we have

$$\frac{d\ln(W/R)}{dN} = \frac{d\ln\omega}{dN} = \frac{1}{\omega}\frac{\partial\omega}{\partial N} > 0,$$

where $\frac{\partial \omega}{\partial N}$ denotes the partial derivative of $\omega(I^*, N, K)$ with respect to N.

Turning next to the case where $I^* = \widetilde{I} < I$, note that the threshold task is given by $\gamma(I^*) = \omega K$. Therefore, $dI^* = \frac{1}{\varepsilon_\gamma} d\ln\omega$ (where recall that ε_γ is the semi-elasticity of the γ function as defined in the proposition). Therefore, $\frac{d\ln(W/R)}{dN} = \frac{d\ln\omega}{dN}$, and we can compute this total derivative as claimed in proposition:

$$\frac{d\ln\omega}{dN} = \frac{1}{\omega}\frac{\partial\omega}{\partial N} + \frac{1}{\omega}\frac{\partial\omega}{\partial I^*}\frac{1}{\varepsilon_\gamma}\frac{d\ln\omega}{dN} = \frac{\frac{1}{\omega}\frac{\partial\omega}{\partial N}}{1 - \frac{1}{\omega}\frac{\partial\omega}{\partial I^*}\frac{1}{\varepsilon_\gamma}} = \frac{\sigma_{\text{cons}} + \varepsilon_L}{\sigma_{\text{free}} + \varepsilon_L}\frac{1}{\omega}\frac{\partial\omega}{\partial N}.$$

To conclude the proposition, we specialize to the case in which Assumption 2 holds. The expressions for the partial derivative $\frac{\partial\omega}{\partial I^*}$, $\frac{\partial\omega}{\partial N}$ and $\hat{\sigma}$ presented in the proposition follow directly from differentiating equation (13) in the main text. Finally, the definition of σ_{free} in the proposition implies that in this case,

$$\sigma_{\text{free}} = (\hat{\sigma} + \varepsilon_L)\left(1 - \frac{1}{\omega}\frac{\partial\omega}{\partial I^*}\frac{1}{\varepsilon_\gamma}\right) - \varepsilon_L = \hat{\sigma} + \frac{1}{\varepsilon_\gamma}\Lambda_I,$$

which proofs the claims in Proposition 2 in the main text. ■

Proof of Proposition 3: The formulas provided for $d\ln Y\mid_{K,L}$ in this proposition hold under Assumption 2, and we impose this assumption in this proof.

We start by deriving the formulas for $d\ln Y\mid_{K,L}$ in the case in which technology binds and $I^* = I < \widetilde{I}$. To do so, we first consider a change in dN and totally differentiate equation (12) in the main text:

$$\begin{aligned}
d\ln Y\mid_{K,L} =& \frac{B}{(1-\eta)Y}\left[\frac{Y(1-\eta)}{B}\right]^{\frac{1}{\hat{\sigma}}}\frac{1}{\hat{\sigma}-1}\left(\gamma(N)^{\hat{\sigma}-1}\left(\frac{\int_{I^*}^{N}\gamma(i)^{\hat{\sigma}-1}di}{L}\right)^{\frac{1-\hat{\sigma}}{\hat{\sigma}}} - \left(\frac{I^*-N+1}{K}\right)^{\frac{1-\hat{\sigma}}{\hat{\sigma}}}\right)dN \\
=& \frac{B}{(1-\eta)Y}\left[\frac{Y(1-\eta)}{B}\right]^{\frac{1}{\hat{\sigma}}}\frac{1}{\hat{\sigma}-1}\left(\gamma(N)^{\hat{\sigma}-1}\left(\frac{B^{1-\hat{\sigma}}W^{\hat{\sigma}}}{(1-\eta)Y}\right)^{\frac{1-\hat{\sigma}}{\hat{\sigma}}} - \left(\frac{B^{1-\hat{\sigma}}R^{\hat{\sigma}}}{(1-\eta)Y}\right)^{\frac{1-\hat{\sigma}}{\hat{\sigma}}}\right)dN \\
=& B^{\hat{\sigma}-1}\frac{1}{1-\hat{\sigma}}\left(R^{1-\hat{\sigma}} - \left(\frac{W}{\gamma(N)}\right)^{1-\hat{\sigma}}\right)dN.
\end{aligned}$$

Likewise, following a change in dI^*, we have

$$\begin{aligned}
d\ln Y\mid_{K,L} =& \frac{B}{(1-\eta)Y}\left[\frac{Y(1-\eta)}{B}\right]^{\frac{1}{\hat{\sigma}}}\frac{1}{\hat{\sigma}-1}\left(\left(\frac{I^*-N+1}{K}\right)^{\frac{1-\hat{\sigma}}{\hat{\sigma}}} - \gamma(I)^{\hat{\sigma}-1}\left(\frac{\int_{I^*}^{N}\gamma(i)^{\hat{\sigma}-1}di}{L}\right)^{\frac{1-\hat{\sigma}}{\hat{\sigma}}}\right)dI \\
=& \frac{B}{(1-\eta)Y}\left[\frac{Y(1-\eta)}{B}\right]^{\frac{1}{\hat{\sigma}}}\frac{1}{\hat{\sigma}-1}\left(\left(\frac{B^{1-\hat{\sigma}}R^{\hat{\sigma}}}{(1-\eta)Y}\right)^{\frac{1-\hat{\sigma}}{\hat{\sigma}}} - \gamma(I)^{\hat{\sigma}-1}\left(\frac{B^{1-\hat{\sigma}}W^{\hat{\sigma}}}{(1-\eta)Y}\right)^{\frac{1-\hat{\sigma}}{\hat{\sigma}}}\right)dI \\
=& B^{\hat{\sigma}-1}\frac{1}{1-\hat{\sigma}}\left(\left(\frac{W}{\gamma(I)}\right)^{1-\hat{\sigma}} - R^{1-\hat{\sigma}}\right)dI.
\end{aligned}$$

We now derive the formulas for the impact of technology on factor prices. Let s_L denote the labor share in net output. Because $WL + RK = (1-\eta)Y$, we obtain

$$s_L d\ln W + (1-s_L)d\ln R = d\ln Y\mid_{K,L}. \tag{B13}$$

Moreover, Proposition 2 implies

$$d\ln W - d\ln R = \frac{1}{\hat{\sigma} + \varepsilon_L}\Lambda_N dN - \frac{1}{\hat{\sigma} + \varepsilon_L}\Lambda_I dI. \tag{B14}$$

Solving the system of equations given by (B13) and (B14), we obtain the formulas for $d\ln W$ and $d\ln R$ in the proposition.

To establish the existence of the threshold $\widetilde{K}$, we substitute $1 - s_L = (I^* - N + 1)B^{\hat{\sigma}-1}R^{1-\hat{\sigma}}$—this is the share of capital in output net of intermediates—in the formula for $\frac{d\ln W}{dI}$ given in the proposition. We find that automation reduces wages if and only if:

$$\frac{1}{1-\hat{\sigma}}\left[\left(\frac{W}{R}\frac{1}{\gamma(I^*)}\right)^{1-\hat{\sigma}} - 1\right] < (I^* - N + 1)\Lambda_I.$$

Let $\underline{K}$ be the level of capital at which $\frac{W}{\gamma(I^*)} = R$. For $K > \underline{K}$, we have that $\frac{W}{\gamma(I^*)} \geq R$, and thus $I^* = I < \widetilde{I}$. At $\underline{K}$, the above inequality holds. Also, the left-hand side of the above inequality is a continuous and increasing function of W/R. This implies that there exists a threshold $\widetilde{K} > \underline{K}$ such that, the above inequality holds for $K \in (\underline{K}, \widetilde{K})$ but is reversed for $K > \widetilde{K}$.

Consider next the case where $I^* = \widetilde{I} < I$. In this case we have:

$$\begin{aligned} d\ln Y\mid_{K,L} &= B^{\hat{\sigma}-1}\frac{1}{1-\hat{\sigma}}\left(R^{1-\hat{\sigma}} - \left(\frac{W}{\gamma(N)}\right)^{1-\hat{\sigma}}\right)dN + B^{\hat{\sigma}-1}\frac{1}{1-\hat{\sigma}}\left(\left(\frac{W}{\gamma(\widetilde{I})}\right)^{1-\hat{\sigma}} - R^{1-\hat{\sigma}}\right)dI^* \\ &= B^{\hat{\sigma}-1}\frac{1}{1-\hat{\sigma}}\left(R^{1-\hat{\sigma}} - \left(\frac{W}{\gamma(N)}\right)^{1-\hat{\sigma}}\right)dN. \end{aligned}$$

Thus, changes in I^* do not affect aggregate output because the marginal firm at $\widetilde{I}$ is indifferent between producing with capital or producing with labor. On the other hand, because I is not binding, changes in I do not affect aggregate output.

We derive the formulas for the impact of technology on factor prices as before, except that equation (B14) now becomes

$$d\ln W - d\ln R = \frac{1}{\sigma_{\text{free}} + \varepsilon_L}\Lambda_N dN.$$

■

Remaining Proofs from Section 3

We start by providing an additional lemma showing that, for a path of technology in which $g(t) = g$ and $n > \max\{\overline{n}, \widetilde{n}(\rho)\}$, the resulting production function $F(k, L; n)$ satisfies the Inada conditions required in a BGP.

Lemma B1 (Inada conditions) *Suppose that Assumptions 1′ and 2 hold. Consider a path of technology in which $n(t) \to n$ and $g(t) \to g$. Let $F(k, L; n)$ denote net output introduced in the proof of Proposition 4. If $\rho \in (\rho_{min}, \rho_{max})$ and $n > \max\{\overline{n}(\rho), \widetilde{n}(\rho)\}$ we have that F satisfies the Inada conditions*

$$\lim_{\phi\to 0} F_K(\phi, 1; n) > \rho + \delta + \theta g \qquad \lim_{\phi\to\infty} F_K(\phi, 1; n) < \rho + \delta + \theta g.$$

Proof. Let $\phi = \frac{k}{L}$. Let $MPK(\phi) = F_K(\phi, 1; n)$ and $w(\phi) = F_L(\phi, 1; n)$ denote the rental rate of capital and the wage at this ratio, respectively.

When $n > \max\{\overline{n}(\rho), \widetilde{n}(\rho)\}$, these factor prices satisfy the system of equations given by the ratio of the market-clearing conditions (A3) and (A4),

$$\phi = \frac{(1-n)c^u(MPK(\phi))^{\zeta-\sigma}MPK(\phi)^{-\zeta}}{\int_0^n \gamma(i)^{\zeta-1}c^u(w(\phi)/\gamma(i))^{\zeta-\sigma}w(\phi)^{-\zeta}},$$

together with the generalized ideal price index condition (A5), which we can rewrite succinctly as:

$$B^{1-\hat{\sigma}} = (1-n)c^u(MPK(\phi))^{1-\sigma} + \int_0^n c^u(w(\phi)/\gamma(i))^{1-\sigma}di. \tag{B15}$$

We start by considering the limit case in which $\phi = 0$. The factor-demand equation requires that either (i) $MPK(\phi) = \infty$, or (ii) $w(\phi) = 0$. In the first case, we have $MPK(\phi) > \rho + \delta + \theta g$ as claimed. In the second case we have:

$$c^u(0) = \begin{cases} 0 & \text{if } \zeta \geq 1 \\ c_0^u & \text{if } \zeta < 1. \end{cases}$$

We show that in both cases $MPK(0) > \rho + \delta + \theta g$:

1. Suppose that $\zeta \geq 1$. For the ideal price index condition in (B15) to hold, we require $\sigma < 1$ (otherwise the right-hand side diverges). Moreover, the ideal price index condition in (B15) implies that $MPK(0)$ is implicitly given by:

   $$(1-n)c^u(MPK(0))^{1-\sigma} = B^{1-\hat{\sigma}}.$$

   First, suppose that $\rho \leq \overline{\rho}$. We have that

   $$c^u(MPK(0))^{1-\sigma} > (1-n)c^u(MPK(0))^{1-\sigma} = B^{1-\hat{\sigma}} = c^u(\overline{\rho} + \delta + \theta g)^{1-\sigma}.$$

   Here we have used the fact that $n > 0$ and the definition of $\overline{\rho}$ introduced in Lemma A2. Because $\sigma < 1$, the above inequality implies $MPK(0) > \overline{\rho} + \delta + \theta g \geq \rho + \delta + \theta g$ as claimed.

   Finally, suppose that $\rho > \overline{\rho}$. Because $n > \overline{n}(\rho)$, we have:

   $$\begin{aligned}(1-\overline{n}(\rho))c^u(MPK(0))^{1-\sigma} &> (1-n)c^u(MPK(0))^{1-\sigma} \\ &= B^{1-\hat{\sigma}} \\ &= (1-\overline{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma} + \int_0^{\overline{n}(\rho)} c^u((\rho+\delta+\theta g)/\gamma(i))^{1-\sigma}di \\ &> (1-\overline{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma}.\end{aligned}$$

   Here we have also used the definition of $\overline{n}(\rho)$ introduced in Lemma A2. Because $\sigma < 1$, the above inequality implies $MPK(0) > \rho+\delta+\theta g$ as claimed (recall that in this region $\overline{n}(\rho) < 1$).

2. Suppose that $\zeta < 1$. We have that $0 < c_0^u < c^u(x)$ for all $x > 0$. The ideal price index condition in (B15) implies that $MPK(0)$ is implicitly given by:

$$(1-n)c^u(MPK(0))^{1-\sigma} + n c_0^{u\,1-\sigma} = B^{1-\hat{\sigma}}.$$

When $\sigma < 1$, we have the following series of inequalities:

$$\begin{aligned}(1-\overline{n}(\rho))c^u(MPK(0))^{1-\sigma} + \overline{n}(\rho)c_0^{u\,1-\sigma} &> (1-n)c^u(MPK(0))^{1-\sigma} + n c_0^{u\,1-\sigma} \\ &= B^{1-\hat{\sigma}} \\ &= (1-\overline{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma} \\ &+ \int_0^{\overline{n}(\rho)} c^u((\rho+\delta+\theta g)/\gamma(i))^{1-\sigma} di \\ &> (1-\overline{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma} + \overline{n}(\rho)c_0^{u\,1-\sigma}.\end{aligned}$$

Here, we have used the fact that $n > \overline{n}(\rho)$ and $0 < c_0^u < c^u(x)$ for all $x > 0$, and the definition of $\overline{n}(\rho)$ introduced in Lemma A2. Because $\sigma < 1$, the above inequality implies $MPK(0) > \rho + \delta + \theta g$ as claimed (recall that in this region $\overline{n}(\rho) < 1$).

When $\sigma > 1$, the previous inequalities are reversed, and thus

$$(1-\overline{n}(\rho))c^u(MPK(0))^{1-\sigma} + \overline{n}(\rho)c_0^{u\,1-\sigma} < (1-\overline{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma} + \overline{n}(\rho)c_0^{u\,1-\sigma}.$$

Because $\sigma > 1$, the above inequality implies $MPK(0) > \rho + \delta + \theta g$ as claimed.

We next consider the limit case in which $\phi = \infty$. With a slight abuse of notation, we define $MPK(\infty) = \lim_{\phi\to\infty} MPK(\phi)$ and $w(\infty) = \lim_{\phi\to\infty} w(\phi)$. The factor-demand equation requires that either (i) $MPK(\infty) = 0$, or (ii) $w(\infty) = \infty$. In the first case, $MPK(\infty) < \rho + \delta + \theta g$. In the second case, we have

$$c^u(\infty) = \begin{cases} \infty & \text{if } \zeta \leq 1 \\ c_\infty^u & \text{if } \zeta > 1. \end{cases}$$

We show that in both cases $MPK(\infty) < \rho + \delta + \theta g$.

1. Suppose that $\zeta \leq 1$. For the ideal price index condition in (B15) to hold, we require $\sigma > 1$ (otherwise the right-hand side diverges). Moreover, the ideal price index condition in (B15) implies that $MPK(\infty)$ is implicitly given by

$$(1-n)c^u(MPK(\infty))^{1-\sigma} = B^{1-\hat{\sigma}}.$$

First, suppose that $\rho \geq \overline{\rho}$. Then

$$c^u(MPK(\infty))^{1-\sigma} > (1-n)c^u(MPK(\infty))^{1-\sigma} = B^{1-\hat{\sigma}} = c^u(\overline{\rho}+\delta+\theta g)^{1-\sigma}.$$

Here we have used the fact that $n > 0$ and the definition of $\overline{\rho}$ introduced in Lemma A2. Because $\sigma > 1$, the above inequality implies $MPK(\infty) < \overline{\rho} + \delta + \theta g \leq \rho + \delta + \theta g$ as claimed.

Finally, suppose that $\rho < \overline{\rho}$. Because $n > \widetilde{n}(\rho)$, we have

$$\begin{aligned}
(1-\widetilde{n}(\rho))c^u(MPK(\infty))^{1-\sigma} &> (1-n)c^u(MPK(\infty))^{1-\sigma} \\
&= B^{1-\hat{\sigma}} \\
&= (1-\widetilde{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma} + \int_0^{\widetilde{n}(\rho)} c^u((\rho+\delta+\theta g)\gamma(i))^{1-\sigma} di \\
&> (1-\widetilde{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma}.
\end{aligned}$$

Here we have also used the definition of $\widetilde{n}(\rho)$ introduced in Lemma A2. Because $\sigma > 1$, the above inequality implies $MPK(\infty) < \rho + \delta + \theta g$ as claimed (recall that in this region $1 > \widetilde{n}(\rho)$).

2. Suppose that $\zeta > 1$. We have that $0 < c^u(x) < c^u_\infty$ for all $x < \infty$. The ideal price index condition in (B15) implies that $MPK(\infty)$ is implicitly given by

$$(1-n)c^u(MPK(\infty))^{1-\sigma} + n{c^u_\infty}^{1-\sigma} = B^{1-\hat{\sigma}}.$$

When $\sigma < 1$, we also have

$$\begin{aligned}
(1-\widetilde{n}(\rho))c^u(MPK(\infty))^{1-\sigma} + \widetilde{n}(\rho){c^u_\infty}^{1-\sigma} &< (1-n)c^u(MPK(\infty))^{1-\sigma} + n{c^u_\infty}^{1-\sigma} \\
&= B^{1-\hat{\sigma}} \\
&= (1-\widetilde{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma} \\
&\quad + \int_0^{\widetilde{n}(\rho)} c^u((\rho+\delta+\theta g)\gamma(i))^{1-\sigma} di \\
&< (1-\widetilde{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma} + \widetilde{n}(\rho){c^u_\infty}^{1-\sigma}.
\end{aligned}$$

Here, we have used the fact that $n > \widetilde{n}(\rho)$ and $0 < c^u(x) < c^u_\infty$ for all $x < \infty$, and the definition of $\widetilde{n}(\rho)$ introduced in Lemma A2. Because $\sigma < 1$, this series of inequalities implies $MPK(\infty) < \rho + \delta + \theta g$ as claimed (recall that in this region $\widetilde{n}(\rho) < 1$).

When $\sigma > 1$, the previous inequalities are reversed, and

$$(1-\widetilde{n}(\rho))c^u(MPK(\infty))^{1-\sigma} + \widetilde{n}(\rho){c^u_\infty}^{1-\sigma} > (1-\widetilde{n}(\rho))c^u(\rho+\delta+\theta g)^{1-\sigma} + \widetilde{n}(\rho){c^u_\infty}^{1-\sigma}.$$

Because $\sigma > 1$, this inequality implies $MPK(\infty) < \rho + \delta + \theta g$, completing the proof.

■

Proof of Global Stability for Part 2 of Proposition 4: Here we provide the details of global stability of the interior equilibrium where all automated tasks are immediately produced with capital (part 2 of Proposition 4). In particular, we show that the BGP given by $k(t) = k_B$, $c(t) = c_B$ and $L(t) = L_B$ is globally stable.

For a given level of capital and consumption, we can define the equilibrium labor supply schedule, $L^E(k,c)$, implicitly as the solution to the first-order condition

$$\nu'(L^E(k,c))e^{\nu(L^E(k,c))\frac{\theta-1}{\theta}} = \frac{F_L(k, L^E(k,c))}{c}.$$

The left-hand side of this equation is increasing in L^E. Thus, the optimal labor supply $L^E(k,c)$ is increasing in k (because of the substitution effect) and is decreasing in c (because of the income effect). In addition, because F_L is homogeneous of degree zero, one can verify that $\frac{L}{k} > L^E_k > 0$, so that labor responds less than one-to-one to an increase in capital.

Any dynamic equilibrium must solve the system of differential equations

$$\begin{aligned}\frac{\dot{c}(t)}{c(t)} =& \frac{1}{\theta}(F_K(k(t), L^E(k(t),c(t)); n) - \delta - \rho) - g \\ \dot{k}(t) =& F(k(t), L^E(k(t),c(t)); n) - (\delta+g)k(t) - c(t)e^{\nu(L^E(k(t),c(t)))\frac{\theta-1}{\theta}},\end{aligned}$$

together with the transversality condition in equation (19).

We analyze this system in the (c,k) space. We always have one of the two cases portrayed in Figure B2; either $\lim_{c\to 0} L^E(k,c) = \overline{L}$ or $\lim_{c\to 0} L^E(k,c) = \infty$.

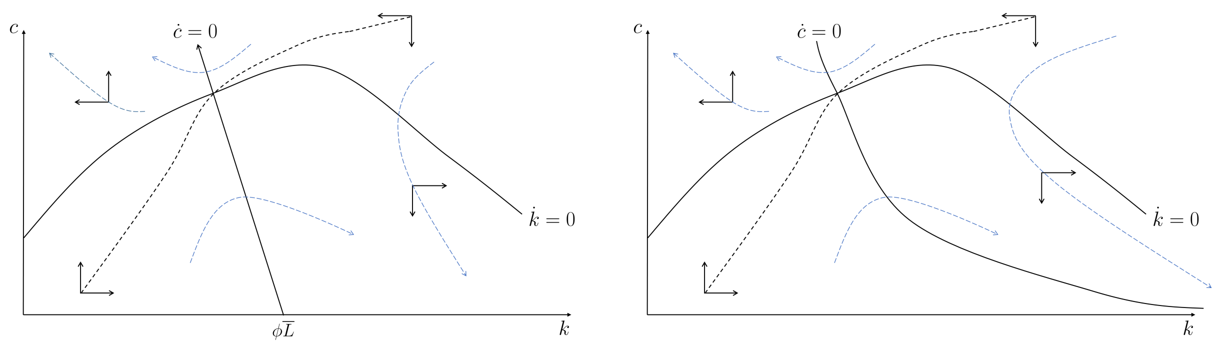

FIGURE B2: The left panel shows the phase diagram of the equilibrium system when $\lim_{c\to 0} L^E(k,c) = \overline{L}$. The right panel shows the phase diagram of the equilibrium system when $\lim_{c\to 0} L^E(k,c) = \infty$.

The locus for $\dot{k} = 0$ yields a curve that defines the maximum level of consumption that can be sustained at each level of capital. This level is determined implicitly by

$$F(k, L^E(k,c); n) - (\delta+g)k = ce^{\nu(L^E(k,c))\frac{\theta-1}{\theta}}.$$

The locus for $\dot{c} = 0$ is given by $k = \phi L^E(k,c)$, which defines a decreasing curve between c and k. Depending on whether $\nu'(L)$ has a vertical asymptote or not, as $c \to 0$, this locus converges to $k = \phi\overline{L}$ (left panel in figure B2) or $k = \infty$ (right panel in figure B2).

Importantly, we always have that, as $c \to 0$, the locus for $\dot{k} = 0$ is above the locus for $\dot{c} = 0$. This is clearly the case when $\lim_{c\to 0} L^E(k,c) = \overline{L}$. To show this when $\lim_{c\to 0} L^E(k,c) = \infty$, consider a point (c_0, k_0) in the locus for $\dot{k} = 0$. We have

$$F_K\left(1, \frac{L^E(k_0,c_0)}{k_0}\right) < F\left(1, \frac{L^E(k_0,c_0)}{k_0}\right) = \delta + g + \mathcal{O}(c_0)$$

Thus, for $c_0 \to 0$, the condition $\rho + (\theta-1)g > 0$ implies

$$F_K\left(1, \frac{L^E(k_0, c_0)}{k_0}\right) < \rho + \delta + \theta g.$$

This inequality implies that $\frac{L^E(k_0,c_0)}{k_0} < \frac{1}{\phi}$, which is equivalent to the point (c_0, k_0) being in the northeast region of the locus for $\dot{c} = 0$.

As shown in Appendix A, both when $\lim_{c\to 0} L^E(k,c) = \overline{L}$ or $\lim_{c\to 0} L^E(k,c) = \infty$, we have a unique interior equilibrium at (c_B, k_B). Moreover, because as $c \to 0$ the locus for $\dot{c} = 0$ is below the locus for $\dot{k} = 0$, we must have that the locus for $\dot{c} = 0$ always cuts the locus for $\dot{k} = 0$ from above at (c_B, k_B). Thus, as shown in the phase diagrams in Figure B2, the unique interior equilibrium at (c_B, k_B) is saddle-path stable.

One could also establish local saddle-path stability as follows. Around the interior BGP, the system of differential equations that determines the equilibrium can be linearized as (suppressing the arguments of the derivatives of the production function)

$$\begin{aligned}
\dot{k}(t) =& \left(\rho + (\theta-1)g + \frac{1}{\theta}F_L L_k^E\right)(k(t) - k_B) + \left(-e^{\nu(L_B)\frac{\theta-1}{\theta}} + \frac{1}{\theta}F_L L_c^E\right)(c(t) - c_B) \\
\dot{c}(t) =& \frac{c_B}{\theta}(F_{KK} + F_{KL}L_k^E)(k(t) - k_B) + \frac{c_B}{\theta}F_{KL}L_c^E(c(t) - c_B).
\end{aligned}$$

The characteristic matrix of the system is therefore given by

$$M_{\text{exog}} = \begin{pmatrix} \rho + (\theta-1)g + \frac{1}{\theta}F_L L_k^E & -e^{\nu(L_B)\frac{\theta-1}{\theta}} + \frac{1}{\theta}F_L L_c^E \\ \frac{c_B}{\theta}(F_{KK} + F_{KL}L_k^E) & \frac{c_B}{\theta}F_{KL}L_c^E \end{pmatrix}.$$

To analyze the properties of this matrix, we will use two facts: (i) $F_L L_k^E + c_B F_{KL} L_c^E = 0$ and (ii) $F_{KK} + F_{KL}L_k^E < 0$. First, (i) follows by implicitly differentiating the optimality condition for labor, which yields:

$$L_k^E = \frac{\frac{1}{c}F_{Lk}}{e^{\nu(L)\frac{\theta-1}{\theta}}\left(\nu''(L) + \frac{\theta-1}{\theta}\nu'^2\right) - \frac{1}{c}F_{LL}} \qquad L_c^E = -\frac{\frac{1}{c^2}F_L}{e^{\nu(L)\frac{\theta-1}{\theta}}\left(\nu''(L) + \frac{\theta-1}{\theta}\nu'^2\right) - \frac{1}{c}F_{LL}}.$$

Next, (ii) follows by noting that, because $L_k^E < \frac{L^E}{k}$, we have

$$F_{KK} + F_{KL}L_k^E < F_{KK} + F_{KL}\frac{L^E}{k} = 0.$$

Using these facts, we can compute the trace of M_{exog} as

$$\text{Tr}(M_{\text{exog}}) = \rho + (\theta-1)g + \frac{1}{\theta}F_L L_k^E + \frac{c_B}{\theta}F_{KL}L_c^E = \rho + (\theta-1)g > 0.$$

In addition, the determinant of M_{exog} is given by:

$$\begin{aligned}
\text{Det}(M_{\text{exog}}) =& \frac{c_B}{\theta}F_{KL}L_c^E\left(\rho + (\theta-1)g + \frac{1}{\theta}F_L L_k^E\right) \\
&- \frac{c_B}{\theta}(F_{KK} + F_{KL}L_k^E)\left(\frac{1}{\theta}F_L L_c^E - e^{\nu(L_B)\frac{\theta-1}{\theta}}\right) < 0
\end{aligned} \tag{B16}$$

The inequality follows by noting that $F_{KL}L_c^E < 0$, $\rho + (\theta-1)g + \frac{1}{\theta}F_L L_k^E > 0$, $F_{KK} + F_{KL}L_k^E < 0$, and $\frac{1}{\theta}F_L L_c^E - e^{\nu(L)\frac{\theta-1}{\theta}} < 0$.

(The negative determinant is equivalent to the fact established above that the curve for $\dot{c} = 0$ cuts the curve for $\dot{k} = 0$ from above. Moreover, the algebra here shows that, at the intersection (c_B, k_B), the locus for $\dot{k}$ is increasing).

The sign of the trace and the determinant imply that the matrix has one positive and real eigenvalue and one negative and real eigenvalue. Theorem 7.19 in Acemoglu (2009) shows that, locally, the economy with exogenous technology is saddle-path stable as wanted.

To show the global stability of the unique BGP (c_B, k_B), we need to rule out two types of paths: the candidate paths that converge to zero capital, which we will show are not feasible, and the candidate paths that converge to zero consumption, which we will show are not optimal.

To rule out the paths that converge to zero capital, note that such paths converge to an allocation with $k(t) = 0$ and $c(t) > \underline{c}$. Here $\underline{c} \geq 0$ is the maximum level of consumption that can be sustained when $k = 0$, which is given by:

$$F(0, L^E(0, \underline{c})) = \underline{c}e^{\nu(L^E(0,\underline{c}))\frac{\theta-1}{\theta}}.$$

To rule out the paths that converge to zero consumption, we show that they violate the transversality condition in equation (19). In all these paths we have $c(t) \to 0$. There are two possible paths for capital. Either capital converges to $\overline{k}$—even at zero consumption the economy only sustains a finite amount of capital—, or capital grows with no bound. In the first case, note that:

$$F_K\left(1, \frac{L^E(k,c)}{k}\right) \leq F\left(1, \frac{L^E(k,c)}{k}\right) = \delta + g.$$

Thus, the transversality condition in (19) does not hold. In the second case, we have that capital grows at an asymptotic rate of $F\left(1, \frac{L^E(k,c)}{k}\right) - \delta - g$. This is greater than or equal to the discount rate used in the transversality condition in equation (19), which is $F_K\left(1, \frac{L^E(k,c)}{k}\right) - \delta - g$. Thus, the transversality condition does not hold in this case either. ■

Proof of Proposition 5: We prove the proposition in the more general case in which Assumption 2′ holds.

Proposition 4 shows that for this path of technology the economy admits a unique BGP.

If $n < \overline{n}(\rho)$, we have that in the BGP $n^*(t) = \overline{n}(\rho) > n$. Thus, small changes in n do not affect the BGP equations; n does not affect effective wages, employment, or the labor share.

If $n > \overline{n}(\rho)$, we have that in the BGP $n^*(t) = n$. In this case, the behavior of the effective wages follows from the formulas for $w_I'(n)$ and $w_N'(N)$ in equation (A8), whose signs can be determined from Lemma A2.

To characterize the behavior of employment, note that we can rewrite the first-order condition for the BGP level of employment in equation (18) as

$$\frac{1}{L\nu'(L)e^{\frac{\theta-1}{\theta}\nu(L(t))}} = \frac{c}{wL} = \frac{1}{s_L}\frac{\rho + (\theta-1)g}{\rho + \delta + \theta g} + \frac{\delta + g}{\rho + \delta + \theta g}.$$

It follows that, asymptotically, there is an increasing relationship between employment and the labor share (recall that the joint concavity of the utility function requires $\nu'(L)e^{\frac{\theta-1}{\theta}\nu(L(t))}$ to increase in L). Thus, the BGP level of employment is given by the increasing function $L^{LR}(\omega)$, whose elasticity we denote by ε_L^{LR}.

To characterize the behavior of the labor share we use Lemma A1. This lemma was derived for the static model when the labor supply was given by $L^s(\omega)$, but we can use it here to describe the asymptotic behavior of the economy when the supply of labor is given by $L^{LR}(\omega)$.

We consider two cases. First, suppose that $\sigma_{\text{const}} \leq 1$. Let $k_I(n)$ denote the BGP value for $K(t)/\gamma(I(t))$. Recall that the function $\omega(I^*, N, K)$ yields the value of $\omega = \frac{W}{RK}$ when the level of technology is given by I^*, N, and the stock of capital is given by K. Thus, the definition of $w_I(n)$ and $k_I(n)$ implies that:

$$\omega(0, n, k_I(n)) = \frac{w_I(n)}{(\rho+\delta+\theta g)k_I(n)}.$$

Differentiating this expression, we obtain

$$k_I'(n) = \frac{w_I'(n)\frac{1}{Rk} - \frac{\partial\omega}{\partial N}}{\frac{\omega}{k}\frac{1+\varepsilon_L^{LR}}{\sigma_{\text{const}}+\varepsilon_L^{LR}}}.$$

Using this expression for $k_I'(n)$, it follows that the total effect of technology on ω is given by

$$\begin{aligned}\frac{d\omega}{dn} =& \frac{\partial\omega}{\partial N} + \frac{\partial\omega}{\partial K}k_I'(n)\\ =& \frac{\partial\omega}{\partial N}\left(\frac{\sigma_{\text{const}}+\varepsilon_L^{LR}}{1+\varepsilon_L^{LR}}\right) + \frac{w_I'(n)}{Rk}\left(\frac{1-\sigma_{\text{const}}}{1+\varepsilon_L^{LR}}\right).\end{aligned}$$

Because $\frac{\partial\omega}{\partial N} > 0$ and $w_I'(n) > 0$, we have that, whenever $\sigma_{\text{const}} \leq 1$, ω is increasing in n. Moreover, because the BGP level of employment is given by the increasing function $L^{LR}(\omega)$, n raises employment too.

Next suppose that $\sigma_{\text{const}} > 1$. Let $k_N(n)$ denote the BGP value for $K(t)/\gamma(N(t))$. Using an analogous reasoning as before, we get

$$\omega(-n, 0, k_N(n)) = \frac{w_N(n)}{(\rho+\delta+\theta g)k_N(n)}.$$

Differentiating this expression, we have

$$k_N'(n) = \frac{w_N'(n)\frac{1}{Rk} + \frac{\partial\omega}{\partial I^*}}{\frac{\omega}{k}\frac{1+\varepsilon_L^{LR}}{\sigma_{\text{const}}+\varepsilon_L^{LR}}} < 0.$$

Using this expression for $k_N'(n)$, it follows that the total effect of technology on ω is given by

$$\begin{aligned}\frac{d\omega}{dn} =& -\frac{\partial\omega}{\partial I^*} + \frac{\partial\omega}{\partial K}k_N'(n)\\ =& -\frac{\partial\omega}{\partial I^*}\left(\frac{\sigma_{\text{ const}}+\varepsilon_L^{LR}}{1+\varepsilon_L^{LR}}\right) + \frac{w_N'(n)}{Rk}\left(\frac{1-\sigma_{\text{const}}}{1+\varepsilon_L^{LR}}\right).\end{aligned}$$

Because $\frac{\partial \omega}{\partial I^*} < 0$ and $w_N'(n) < 0$, we have that, whenever $\sigma_{\text{const}} \geq 1$, ω is increasing in n. Moreover, because the BGP level of employment is given by the increasing function $L^{LR}(\omega)$, n raises employment too.

The previous observations show that automation reduces the labor share in the long run. In addition, we have shown that $k_N'(n) < 0$, which implies that in response to automation, capital increases above its trend. The induced capital accumulation implies that the impact of automation on the labor share worsens over time if $\sigma_{\text{const}} > 1$ and eases if $\sigma_{\text{const}} < 1$.

When Assumption 2 holds the capital share is given by $(1-n)\left(\frac{R(t)}{B}\right)^{1-\hat{\sigma}}$. In this case, n reduces the capital share and thus increases the labor share. This expression also shows that when the rental rate returns to its BGP level, the induced capital accumulation will cause a further decline in the labor share when $\hat{\sigma} > 1$ and a partial offset when $\hat{\sigma} < 1$. ■

Remaining Proofs from Section 4

Proof of Lemma A3: In a BGP we have that the economy grows at the rate $g = A\frac{\kappa_I \kappa_N}{\kappa_I + \kappa_N} S$.

Suppose that $n \geq \max\{\overline{n}, \widetilde{n}\}$. In this case, we can write the value functions in the BGP as

$$v_N(n) = b\int_0^\infty e^{-(\rho-(1-\theta)g)\tau}\left[c^u\left(w_N(n)e^{g\tau}\right)^{\zeta-\sigma} - c^u\left(\rho+\delta+\theta g\right)^{\zeta-\sigma}\right]d\tau,$$
$$v_I(n) = b\int_0^\infty e^{-(\rho-(1-\theta)g)\tau}\left[c^u\left(\rho+\delta+\theta g\right)^{\zeta-\sigma} - c^u\left(w_I(n)e^{g\tau}\right)^{\zeta-\sigma}\right]d\tau.$$

Thus, the value functions only depend on the unit cost of labor $w_N(n)$ and $w_I(n)$, and on the rental rate, which is equal to $\rho+\delta+\theta g$ in the BGP.

Now consider Taylor expansions of both of these expressions (which are continuously differentiable) around $S = 0$—so that the growth rate of the economy is small. Thus,

$$v_N(n) = \frac{b}{\rho}\left[c^u\left(w_N(n)\right)^{\zeta-\sigma} - c^u\left(\rho+\delta+\theta g\right)^{\zeta-\sigma}\right] + \mathcal{O}(g), \tag{B17}$$
$$v_I(n) = \frac{b}{\rho}\left[c^u\left(\rho+\delta+\theta g\right)^{\zeta-\sigma} - c^u\left(w_I(n)\right)^{\zeta-\sigma}\right] + \mathcal{O}(g).$$

Because $\mathcal{O}(g) \to 0$ as $S \to 0$, we can approximate the above integrals when S is small with the explicit expressions evaluated at $g = 0$.

Differentiating the value functions in (B17) establishes that they are both strictly increasing in n. This follows from the result established in Proposition 5 that, in this region, $w_I(n)$ increases in n and $w_N(n)$ decreases in n. Moreover, as $S \to 0$, both $v_N(n)$ and $v_I(n)$ are positive. Thus, there exists $\widetilde{S}_1$ such that for $S < \widetilde{S}_1$, both $v_N(n)$ and $v_I(n)$ are positive ans strictly increasing in n.

Now suppose that $n \leq \overline{n}(\rho)$ (this case requires that $\rho > \overline{\rho}$). In this region we have $n^*(t) = \overline{n}$. Therefore, newly automated tasks are not immediately produced with capital, and thus

$$v_I(n) = b\int_0^\infty e^{-(\rho-(1-\theta)g)\tau}\left[c^u\left(\min\left\{\rho+\delta+\theta g, \frac{w_I(\overline{n}(\rho))}{\gamma(\overline{n}(\rho)-n)}e^{g\tau}\right\}\right)^{\zeta-\sigma} - c^u\left(\frac{w_I(\overline{n}(\rho))}{\gamma(\overline{n}(\rho)-n)}e^{g\tau}\right)^{\zeta-\sigma}\right]d\tau.$$

The min operator $\min\left\{\rho+\delta+\theta g, \frac{w_I(\overline{n}(\rho))}{\gamma(\overline{n}(\rho)-n)}e^{g\tau}\right\}$ captures the fact that a task that is automated at time t will only generate profits in the future starting at a time $\tau > t$ such that $I^*(\tau) = I(t)$. At this point in time, $\frac{w_I(\overline{n}(\rho))}{\gamma(\overline{n}(\rho)-n)}e^{g\tau} = \rho+\delta+\theta g$, and it becomes profitable to use capital to produce the automated task.

In addition, $w_I(\overline{n}(\rho)) = \rho+\delta+\theta g$. Thus, $\lim_{g\to 0} v_I(n) = 0$ and we have $v_I(n) = \mathcal{O}(g)$ for all $n \leq \overline{n}(\rho)$. On the other hand $v_N(n)$ remains bounded away from zero as $S \to 0$. Thus, there exists $\widetilde{S}_2 > 0$ such that for $S < \widetilde{S}_2$, $\kappa_N v_N(\overline{n}) > \kappa_I v_I(\overline{n}) > 0$ as claimed, and $v_I(\overline{n}) = \mathcal{O}(g)$.

Finally, consider the case where $n < \widetilde{n}(\rho)$ (this case requires that $\rho < \overline{\rho}$). Because $w_N(n)e^{g\tau} > \rho+\delta+\theta g$ and $w_I(n)e^{g\tau} > \rho+\delta+\theta g$ for all $\tau \geq 0$, it follows that, in this region, $v_I(n) > 0 > v_N(n)$ as claimed. Moreover, the derivatives for $w_I(n)$ and $w_N(n)$ in equation (A8) imply that, in this region, both $w_I(n)$ and $w_N(n)$ are decreasing in n. Thus, in this region, $v_I(n)$ is decreasing and $v_N(n)$ is increasing in n.

To complete the proof of this lemma, we simply take $\widetilde{S} = \min\{\widetilde{S}_1, \widetilde{S}_2\}$ if $\theta \geq 1$, and $\widetilde{S} = \min\{\widetilde{S}_1, \widetilde{S}_2, \frac{\rho(\kappa_I+\kappa_N)}{(1-\theta)A\kappa_I\kappa_N}\}$ if $\theta < 1$. This choice also ensures that $\rho + (\theta-1)g > 0$ as required in the Lemma. ■

Proof of local stability for the unique BGP when $\theta > 0$: The local stability analysis applies to the case where $\rho > \bar{\rho}$ and $S < \min\{\widetilde{S}, \hat{S}\}$ and $\frac{\kappa_I}{\kappa_N} > \overline{\kappa}$. In this case, the economy admits a unique BGP.

When $\rho > \bar{\rho}$ and $S < \min\{\widetilde{S}, \hat{S}\}$, we can simplify the characterization of equilibrium. In particular, in this case, starting with initial conditions $n(0) \geq 0$ and $k(0) > 0$, the equilibrium with endogenous technology can be summarized by paths for $\{c(t), k(t), n(t), v(t), S_I(t)\}$ such that:

1. The normalized consumption satisfies the Euler equation:
$$\frac{\dot{c}}{c} = \frac{1}{\theta}(F_K(k, L; n) - \delta - \rho) + \mathcal{O}(g).$$

2. The endogenous labor supply is given by $L^E(k, c; n)$, and is defined implicitly by:
$$c\nu'^{\nu(L)\frac{\theta-1}{\theta}} \geq F_L(k, L; n),$$
with equality if $L^E(k, c; n) > 0$.

3. The capital stock satisfies the resource constraint:
$$\dot{k} = F(k, L; n) + X(k, L; n) - \delta k - ce^{\nu(L)\frac{\theta-1}{\theta}} + \mathcal{O}(g).$$
Here, $X(k, L; n) = b(1-n^*)yc^u(F_K)^{\zeta-\sigma} + by\int_0^{n^*} c^u\left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma} di$ are the profits from the intermediate sales.

4. The transversality condition
$$\lim_{t\to\infty}(k(t) + \pi(t))e^{-\int_0^t F_K(k(s),L(s);n(s))-\delta-\mathcal{O}(g))ds} = 0$$
holds, where $\pi(t) = I(t)v_I(t) + N(t)v_N(t)$ are (the normalized) corporate profits.

5. Technology evolves endogenously according to:

$$\dot{n} = \kappa_N S - (\kappa_I + \kappa_N) G(v) S.$$

6. The value function, $v = \kappa_I v_I - \kappa_N v_N$, satisfies

$$(F_K - \delta - g)v - \dot{v} = b\kappa_I \pi_I(k, L; n) - b\kappa_N \pi_N(k, L; n) + \mathcal{O}(g). \tag{B18}$$

Around $g = 0$, the above system of differential equations is Lipschitz continuous (on their right-hand side, the equations for $\dot{c}$, $\dot{k}$, $\dot{n}$ and $\dot{v}$ have bounded derivatives around the BGP $\{c_B, k_B, n_B, v_B\}$; this can be seen from the matrix containing these derivatives M_{endog}, which we present below). Thus, from the theorem of the continuous dependence of trajectories of a dynamical system on parameters (e.g., Walter, 1998, page 146, Theorem VI), there exists a neighborhood of $g = 0$ and a threshold $\overline{S}_1$ such that for $S < \overline{S}_1$, the trajectories that solve the above system have the same direction as the trajectories of the system evaluated at $g = 0$. In particular, for $S < \overline{S}_1$, the BGP is locally saddle-path stable if and only if it is also locally saddle path stable in the limit in which $g = 0$.

The previous argument shows that, to analyze the local stability of the BGP when $S < \overline{S}_1$, it is sufficient to analyze the limit case in which $g = 0$. In what follows we focus on this limit.

As in the proof of Proposition 4, the Euler equation and the resource constraint can be linearized around the BGP (denoted with the subscript B) as follows:

$$\begin{aligned}
\dot{c} =& \frac{c_B}{\theta}(F_{Kn} + F_{KL}L_n^E)(n - n_B) + \frac{c_B}{\theta}F_{KL}L_c^E(c - c_B) + \frac{c_B}{\theta}(F_{KK} + F_{KL}L_k^E)(k - k_B) \\
\dot{k} =& \left(F_n + \frac{1}{\theta}F_L L_n^E + X_n + X_L L_n^E\right)(n - n_B) + \left(\frac{1}{\theta}F_L L_c^E - e^{\nu(L)\frac{\theta-1}{\theta}} + X_L L_c^E\right)(c - c_B) \\
&+ \left(F_K + \frac{1}{\theta}F_L L_k^E - \delta + X_k + X_L L_k^E\right)(k - k_B).
\end{aligned}$$

Let us denote by X_k, X_L and X_n the partial derivatives of the function $X(k, L; n)$ with respect to each of its arguments.

We now show that, under Assumption 4 (which requires $\sigma > \zeta$), we have $X_k > 0$ and $X_L > 0$—that is, the demand for intermediates is increasing in K and L (which also implies that capital and labor are complements to intermediates). We first show this for X_L. Let us rewrite X as

$$X = \frac{b}{B^{\hat{\sigma}-1}(1-\eta)}kR^{\zeta} + by\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma} di,$$

which implies

$$\begin{aligned}
X_L =& \frac{b}{B^{\hat{\sigma}-1}(1-\eta)}k\zeta R^{\zeta-1}F_{KL} \\
&+ by_L\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma} di + (\zeta - \sigma)by\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma} \varsigma_L(i)\frac{F_{LL}}{F_L}di > 0.
\end{aligned}$$

Here, $\varsigma_L(i)$ is the share of labor in the production of task i. The above inequality follows from the fact that $y_L > 0$, $F_{KL} > 0$, and $F_{LL} < 0$.

We now show that $X_k > 0$. Differentiating the labor market-clearing condition yields

$$\begin{aligned}
\frac{y_k}{y} =&(\sigma-\zeta)\frac{\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma}\gamma(i)^{\zeta-1}\varsigma_L(i)\frac{F_{KL}}{F_L}di}{\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma}\gamma(i)^{\zeta-1}di} + \zeta\frac{F_{KL}}{F_L} + \frac{L_k^E}{L^E} \\
>&(\sigma-\zeta)\frac{\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma}\gamma(i)^{\zeta-1}\varsigma_L(i)\frac{F_{KL}}{F_L}di}{\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma}\gamma(i)^{\zeta-1}di}.
\end{aligned}$$

An application of Chebyshev's inequality then implies

$$\frac{y_k}{y} > (\sigma-\zeta)\frac{\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma}\varsigma_L(i)\frac{F_{KL}}{F_L}di}{\int_0^n c^u \left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma}di}. \tag{B19}$$

(Chebyshev's inequality applies because when $\zeta > 1$, both $\gamma(i)^{\zeta-1}$ and $\varsigma_L(i)$ are increasing in i, and when $\zeta < 1$, both are decreasing in i.)

Therefore,

$$\begin{aligned}
X_k =&b(1-n)y_k c^u(F_K)^{\zeta-\sigma} + (\zeta-\sigma)b(1-n^*)yc^u(F_K)^{\zeta-\sigma}\varsigma_K\frac{F_{KK}}{F_K} \\
&+ by_k\int_0^n c^u\left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma}di + by(\zeta-\sigma)\int_0^n c^u\left(\frac{F_L}{\gamma(i)}\right)^{\zeta-\sigma}\varsigma_L(i)\frac{F_{KL}}{F_L}di > 0,
\end{aligned}$$

where ς_K denotes the share of capital in the production of automated tasks. The inequality then follows from the fact that $y_k > 0$ and $F_{KK} < 0$ (recall that $\sigma > \zeta$ under Assumption 4), and the inequality in equation (B19) derived above.

(When Assumption 2 holds, $X = \frac{\eta}{1-\eta}F(k, L; n)$, and it is clear that $X_k > 0$ and $X_L > 0$).

Let Q_n, Q_k and $Q_c > 0$ denote the derivatives of the right-hand side of (B18) with respect to n, k and c. We can then write the Jacobian of the system of differential equations in terms of the derivatives $\{Q_n, Q_k, Q_c\}$, the derivatives $\{X_n, X_k, X_L\}$, and the derivatives of the function F as follows:

$$\begin{pmatrix}
0 & -(\kappa_I+\kappa_N)G'(0)S & 0 & 0 \\
-Q_n & \rho & -Q_c & -Q_k \\
\frac{c^*}{\theta}(F_{Kn}+F_{KL}L_n^E) & 0 & \frac{c^*}{\theta}F_{KL}L_c^E & \frac{c^*}{\theta}(F_{KK}+F_{KL}L_k^E) \\
F_n+\frac{1}{\theta}F_L L_n^E + X_n + X_L L_n^E & 0 & \frac{1}{\theta}F_L L_c^E - e^{\nu(L)\frac{\theta-1}{\theta}} + X_L L_c^E & \rho + \frac{1}{\theta}F_L L_k^E + X_k + X_L L_k^E
\end{pmatrix}.$$

Denote this matrix by M_{endog}, and its eigenvalues by $\lambda_1, \lambda_2, \lambda_3$, and λ_4. These eigenvalues satisfy the following properties:

- The trace satisfies

$$\text{Tr}(M_{\text{endog}}) = \lambda_1+\lambda_2+\lambda_3+\lambda_4 = 2\rho + X_k + X_L L_k^E > 0.$$

  The last inequality follows from the fact that $X_L > 0$ and $L_k^E > 0$

- The determinant satisfies

$$
\begin{aligned}
\mathrm{Det}(M_{\text{endog}}) =& (\kappa_I+\kappa_N)G'(0)S\times \\
&\left[\frac{c_B}{\theta}\left(\frac{1}{\theta}F_L L_c^E - e^{\nu(L)\frac{\theta-1}{\theta}} + X_L L_c^E\right)\left(Q_n(F_{KK}+F_{KL}L_k^E) - Q_k(F_{Kn}+F_{KL}L_n^E)\right)\right. \\
&+\frac{c_B}{\theta}\left(\rho+\frac{1}{\theta}F_L L_k^E + X_k + X_L L_k^E\right)\left(Q_c(F_{Kn}+F_{KL}L_n^E) - Q_n F_{KL}L_c^E\right) \\
&\left.+\frac{c_B}{\theta}\left(F_n+\frac{1}{\theta}F_L L_n^E + X_n + X_L L_n^E\right)\left(Q_k F_{KL}L_c^E - Q_c(F_{KK}+F_{KL}L_k^E)\right)\right].
\end{aligned}
$$

The expression for the determinant can be further simplified by noting that $Q_c(F_{KK} + F_{KL}L_k^E) = Q_k F_{KL} L_c^E$. To show this, note that the impact of k, c on Q—the relative incentives for automation—depends on the ratio $k/L^E(k,c;n)$. For a given value of n, this ratio determines factor prices and hence Q. Let $\phi = \frac{k}{L^E(k,c;n)}$. Then

$$
Q_k = Q_\phi\left(\frac{1}{L} - \frac{k}{L^2}L_k^E\right) Q_c = -Q_\phi\frac{k}{L^2}L_c^E.
$$

These equations then imply

$$
Q_c = -Q_k\frac{kL_c^E}{L-kL_k^E} = Q_k F_{KL}L_c^E\frac{1}{\frac{kL_k^E-L}{k}F_{KL}} = Q_k F_{KL}L_c^E\frac{1}{F_{KK}+F_{KL}L_k^E},
$$

which gives the desired identity

Replacing this expression for Q_c in the determinant, we get

$$
\begin{aligned}
\mathrm{Det}(M_{\text{endog}}) =& (\kappa_I+\kappa_N)G'(0)S\times \\
&\left[\frac{c_B}{\theta}\left(\frac{1}{\theta}F_L L_c^E - e^{\nu(L)\frac{\theta-1}{\theta}} + X_L L_c^E\right)\left(Q_n(F_{KK}+F_{KL}L_k^E) - Q_k(F_{Kn}+F_{KL}L_n^E)\right)\right. \\
&\left.+\frac{c_B}{\theta}F_{KL}L_c^E\left(\rho+\frac{1}{\theta}F_L L_k^E + X_k + X_L L_k^E\right)\left(Q_k\frac{F_{Kn}+F_{KL}L_n^E}{F_{KK}+F_{KL}L_k^E} - Q_n\right)\right].
\end{aligned}
$$

Because $\kappa_I v_I(n)$ cuts (i.e., is steeper than) $\kappa_N v_N(n)$ from below, we have

$$
Q_n - Q_k\frac{F_{Kn}+F_{KL}L_n^E}{F_{KK}+F_{KL}L_k^E} > 0.
$$

(Note that this expression is equivalent to the derivative of the profit function Q with respect to n when the capital adjusts to keep the interest rate constant. This derivative is positive when $\kappa_I v_I(n)$ cuts $\kappa_N v_N(n)$ from below). Because $F_{KK} + F_{KL}L_k^E < 0$ (as shown in the proof of Proposition 4), we also have that

$$
Q_n(F_{KK}+F_{KL}L_k^E) - Q_k(F_{Kn}+F_{KL}L_n^E) < 0.
$$

Thus, both terms in the determinant are positive and $\mathrm{Det}(M_{\text{endog}}) > 0$ (recall that $L_c^E < 0$ and $L_k^E > 0$, and $X_L, X_k > 0$). Because $\mathrm{Det}(M_{\text{endog}}) = \lambda_1\lambda_2\lambda_3\lambda_4$, this implies that the four eigenvalues have either zero, two or four negative real parts.

Let us define $Z(M_{\text{endog}}) = \lambda_1\lambda_2\lambda_3 + \lambda_2\lambda_3\lambda_4 + \lambda_3\lambda_4\lambda_1 + \lambda_4\lambda_1\lambda_2$. We also have

$$Z(M_{\text{endog}}) = \rho\,\mathrm{Det}(M_{\text{exog}}) + \mathcal{O}(S).$$

From equation (B16), we know $\mathrm{Det}(M_{\text{exog}}) < 0$ and this determinant does not depend on S. Thus, there exists $\overline{S}_2 > 0$ such that for $S < \overline{S}_2$, $Z(M_{\text{endog}}) < 0$. This implies that we cannot have four eigenvalues with positive real parts.[35]

But we also have that $\mathrm{Tr}(M_{\text{endog}}) > 0$ as shown above, and thus not all four eigenvalues can have negative real parts.

These observations show that when $S < \overline{S}_2$, M_{endog} has exactly two eigenvalues with negative real parts. Theorem 7.19 in Acemoglu (2009) shows that, in the limit case in which $g \to 0$, the economy with endogenous technology is saddle-path stable. Let $\check{S} = \min\{\overline{S}_1, \overline{S}_2\}$. Thus, for $S < \check{S}$, the unique BGP is locally stable. ■

Proofs from Section 5

Proof of Proposition 7: We prove this proposition under Assumption 2. Consider an exogenous path for technology in which $\dot{N} = \dot{I} = \Delta$ (with $\rho + (\theta - 1)A_H\Delta > 0$) and suppose that $n(t) > \max\{\overline{n}(\rho), \widetilde{n}(\rho)\}$. This implies that in any candidate BGP $I^*(t) = I(t)$ and $n^*(t) = n(t)$.

Define $M \in [I, N]$ as in the main text. Two equations determine M. First, because firms are indifferent between producing task M with low-skill or high-skill workers, we have

$$\frac{W_H(t)}{W_L(t)} = \frac{\gamma_H(M(t))}{\gamma_L(M(t), t)} = \frac{\gamma_H(M(t))^{1-\xi}}{\Gamma(t - T(M(t)))}.$$

In addition, the relative demand for high-skill and low-skill labor yields

$$\frac{L}{H}\frac{\int_{M(t)}^{N(t)} \gamma_H(i)^{\hat{\sigma}-1}di}{\int_{I(t)}^{M(t)} \gamma_L(i,t)^{\hat{\sigma}-1}di} = \left(\frac{W_H(t)}{W_L(t)}\right)^{\hat{\sigma}}.$$

Combining these two equations, we obtain the equilibrium condition

$$\frac{L}{H}\frac{\int_{M(t)}^{N(t)} \gamma_H(i)^{\hat{\sigma}-1}di}{\int_{I(t)}^{M(t)} \gamma_L(i,t)^{\hat{\sigma}-1}di} = \left(\frac{\gamma_H(M(t))^{1-\xi}}{\Gamma(t - T(M(t)))}\right)^{\hat{\sigma}}.$$

Let $m(t) = M(t) - I(t)$ and $n = N(t) - I(t)$. Using the formula for $\gamma_L(i,t)$ and the change of variables $i = N - i'$ to rewrite the integrals in the previous equations we get that $m(t)$ is uniquely

[35] This is independent of whether these eigenvalues are real or complex. For example, if we had two positive real eigenvalues, $\lambda_1, \lambda_2 > 0$, and a conjugate pair of complex eigenvalues with positive real part, λ and $\overline{\lambda}$, then

$$Z(M_{\text{endog}}) = 2\Re(\lambda)\lambda_1\lambda_2 + |\lambda|^2(\lambda_1 + \lambda_2),$$

which cannot be negative. If we had two conjugate pairs of eigenvalues with positive real parts, $\lambda_1, \overline{\lambda}_1$ and $\lambda_2, \overline{\lambda}_2$, then

$$Z(M_{\text{endog}}) = 2\Re(\lambda_1)|\lambda_2|^2 + 2\Re(\lambda_2)|\lambda_1|^2,$$

which again cannot be negative.

pinned down by

$$\frac{L}{H}\frac{\int_0^{n-m(t)}\gamma_H(i)^{1-\hat{\sigma}}di}{\int_{n-m(t)}^{n}\gamma_H(i)^{\xi(1-\hat{\sigma})}\Gamma\left(\frac{i}{\Delta}\right)^{\hat{\sigma}-1}di}=\frac{\gamma_H(N(t))^{1-\xi}}{\gamma_H(n-m(t))^{\sigma(1-\xi)}\Gamma\left(\frac{n-m(t)}{\Delta}\right)^{\hat{\sigma}}}. \tag{B20}$$

This expression also uses the fact that, because both $\dot{N}=\dot{I}=\Delta$, we have $t-T(i)=\frac{N(t)-i}{\Delta}$. The left-hand side of equation (B20)—the relative demand curve—is decreasing in $m(t)$, converges to zero as $m(t)\to n$, and converges to infinity as $m(t)\to 0$. Moreover, the right-hand side—the comparative advantage schedule—is increasing in $m(t)$. Thus, this equation uniquely determines $m(t)$ as a function of $N(t)$ and n.

To prove the first part of the proposition, consider the case in which $\xi<1$. Taking the limit as $t\to\infty$, we have that the right-hand side of equation (B20) converges to infinity. To maintain the equality, we must have $m(t)\to 0$, which implies that asymptotically $M(t)=I(t)$ and no tasks are allocated to low-skill workers. Moreover, we have that inequality grows without bound, since

$$\frac{W_H(t)}{W_L(t)}\to\frac{\gamma_H(N(t))^{1-\xi}}{\gamma(n)^{1-\xi}\Gamma\left(\frac{n}{\Delta}\right)}\to\infty.$$

To prove the second part of the proposition, consider the case where $\xi=1$. We now show that there is a BGP in which $m(t)=m$ and $\frac{W_H(t)}{W_L(t)}$ is constant. Equation (B20) shows that, in this case, m only depends on n as claimed. Moreover, the wage gap is also constant over time and given by

$$\frac{W_H(t)}{W_L(t)}=\frac{1}{\Gamma\left(\frac{n-m}{\Delta}\right)}.$$

Now, consider an increase in n, and let $s=n-m$ denote the measure of tasks performed by high-skill workers. Holding s constant, the left-hand side of equation (B20) is decreasing in n. Because the left-hand side of equation (B20) is increasing in s and its right-hand side is decreasing in s, we must have that s is also increasing in n. This implies that, as stated in the proposition, the wage gap, which is a decreasing function of s, declines with n. ■

Proof of Proposition 8: We prove this result under the more general Assumption 2′.

From the Bellman equations provided in the main text, it follows that along a BGP we have

$$v_N(n)=b\int_0^{\frac{n}{\Delta}}e^{-(\rho-(1-\theta)g)\tau}c^u\left(w_N(n)e^{g\tau}\right)^{\zeta-\sigma}d\tau,$$
$$v_I(n)=b\int_0^{\frac{1-n}{\Delta}}e^{-(\rho-(1-\theta)g)\tau}c^u\left(\rho+\delta+\theta g\right)^{\zeta-\sigma}d\tau.$$

Here $\Delta=\frac{\kappa_I\kappa_N\iota(n^D)}{\kappa_I\iota(n^D)+\kappa_N}S$ is the endogenous rate at which both technologies grow in a BGP and $g=A\Delta$. As before, a BGP requires that n satisfies

$$\kappa_I\iota(n)v_I(n)=\kappa_N v_N(n).$$

Using these formulas, the proof of the proposition follows from the properties of the effective wages derived in Proposition 5. Following the same steps as in the proof of Proposition 6, we also obtain that the equilibrium in this case is locally stable whenever $\kappa_I \iota(n) v_I(n)$ cuts $\kappa_N v_N(n)$ from below. ■

We now turn to Proposition 9. We prove a similar statement in the more general case in which Assumption $2'$ holds. In particular, we show that:

Proposition B2 (Welfare implications of automation in the general model) *Consider the static economy and suppose that Assumptions 1, $2'$ and 3 hold, and that $I^* = I < \widetilde{I}$. Let $\mathcal{W} = u(C, L)$ denote the welfare of households and let $F(K, L; I, N)$ denote the net output when the amount of labor supplied is L and capital is K.*

1. *Consider the baseline model without labor market frictions, so that the representative household chooses the amount of labor without constraints, and thus $\frac{W}{C} = \nu'(L)$. Then:*

$$\frac{d\mathcal{W}}{dI} = \left(Ce^{-\nu(L)}\right)^{1-\theta} \frac{F_I}{F} > 0,$$
$$\frac{d\mathcal{W}}{dN} = \left(Ce^{-\nu(L)}\right)^{1-\theta} \frac{F_N}{F} > 0.$$

2. *Suppose that there are labor market frictions, so that employment is constrained by a quasi-labor supply curve $L \leq L_{qs}(\omega)$. Suppose also that the quasi-labor supply schedule $L_{qs}(\omega)$ is increasing in ω, has an elasticity $\widetilde{\varepsilon}_L > 0$, and is binding in the sense that $\frac{W}{C} > \nu'(L)$. Then:*

$$\frac{d\mathcal{W}}{dI} = \left(Ce^{-\nu(L)}\right)^{1-\theta} \left[\frac{F_I}{F} + L\left(\frac{W}{C} - \nu'(L)\right) \frac{\widetilde{\varepsilon}_L}{\omega} \frac{\partial \omega}{\partial I^*}\right] \lesseqgtr 0.$$
$$\frac{d\mathcal{W}}{dN} = \left(Ce^{-\nu(L)}\right)^{1-\theta} \left[\frac{F_N}{F} + L\left(\frac{W}{C} - \nu'(L)\right) \frac{\widetilde{\varepsilon}_L}{\omega} \frac{\partial \omega}{\partial N}\right] > 0.$$

Proof. The unconstrained allocation of employment solves

$$\mathcal{W} = \max_{L \geq 0} u(F(K, L; I, N), L).$$

Thus, the envelope theorem implies

$$\mathcal{W}_I = u_C F_I = \left(Ce^{-\nu(L)}\right)^{1-\theta} \frac{F_I}{F} > 0$$

(recall that $F_I > 0$ because we assumed $I^* = I$) and also

$$\mathcal{W}_N = u_C F_N = \left(Ce^{-\nu(L)}\right)^{1-\theta} \frac{F_N}{F} > 0$$

(recall that $F_N > 0$ because we imposed Assumption 3).

Now suppose that $L \leq L_{qs}(\omega)$. The allocation of employment now solves:

$$\mathcal{W} = \max_{L \geq 0} u(F(K, L; I, N), L) + \lambda(L_{qs}(\omega) - L),$$

where $\lambda = u_c F_L + u_L = cu_c \left(\frac{F_L}{c} - \nu'(L)\right) > 0$ is the multiplier on the employment constraint (by assumption this constraint is binding). Using the envelope theorem,

$$\mathcal{W}_I = u_C F_I + \lambda L'_{qs}(\omega) \frac{\partial \omega}{\partial I^*} = \left(Ce^{-\nu(L)}\right)^{1-\theta} \left[\frac{F_I}{F} + L\left(\frac{W}{C} - \nu'(L)\right) \frac{\widetilde{\varepsilon}_L}{\omega} \frac{\partial \omega}{\partial I^*}\right] \lesseqgtr 0,$$

and

$$\mathcal{W}_N = u_C F_N + \lambda L'_{qs}(\omega)\frac{\partial \omega}{\partial N} = \left(Ce^{-\nu(L)}\right)^{1-\theta}\left[\frac{F_N}{F} + L\left(\frac{W}{C} - \nu'(L)\right)\frac{\widetilde{\varepsilon}_L}{\omega}\frac{\partial \omega}{\partial N}\right] > 0.$$

The expressions presented in Proposition 9 follow from the previous two equations because when Assumption 2 holds, we also have

$$\begin{aligned}\frac{F_I}{F} =& \frac{B^{\hat{\sigma}-1}}{1-\hat{\sigma}}\left(\left(\frac{W}{\gamma(I)}\right)^{1-\hat{\sigma}} - R^{1-\hat{\sigma}}\right) \qquad & \frac{\widetilde{\varepsilon}_L}{\omega}\frac{\partial \omega}{\partial I} =& -\frac{\widetilde{\varepsilon}_L}{\hat{\sigma}+\widetilde{\varepsilon}_L}\Lambda_I \\ \frac{F_N}{F} =& \frac{B^{\hat{\sigma}-1}}{1-\hat{\sigma}}\left(R^{1-\hat{\sigma}} - \left(\frac{W}{\gamma(N)}\right)^{1-\hat{\sigma}}\right) \qquad & \frac{\widetilde{\varepsilon}_L}{\omega}\frac{\partial \omega}{\partial N} =& \frac{\widetilde{\varepsilon}_L}{\hat{\sigma}+\widetilde{\varepsilon}_L}\Lambda_N.\end{aligned}$$

■

Properties of the constraint efficient allocation:

We now derive the constrained efficient allocation both when the labor market is frictionless and when there is a friction as the one introduced in Proposition 9. We focus on the case in which Assumption 2 holds, although similar insights apply in general.

First the planner removes markups. This implies that net output is given by

$$\begin{aligned}F^p(K,L;I^*,N) =& \mu^{\frac{\eta}{\eta-1}}B\left[(I^*-N+1)^{\frac{1}{\hat{\sigma}}}K^{\frac{\hat{\sigma}-1}{\hat{\sigma}}} + \left(\int_{I^*}^{N}\gamma(i)^{\hat{\sigma}-1}di\right)^{\frac{1}{\hat{\sigma}}}L^{\frac{\hat{\sigma}-1}{\hat{\sigma}}}\right]^{\frac{\hat{\sigma}}{\hat{\sigma}-1}}, \\ =& \mu^{\frac{\eta}{\eta-1}}F(K,L;I^*,N).\end{aligned}$$

Using this expression, we can write the planner's problem as:

$$\max_{C(t),L(t),S_I(t),S_N(t)} \int_0^\infty e^{-\rho t}\frac{[C(t)e^{-\nu(L(t))}]^{1-\theta}-1}{1-\theta}dt$$

Subject to

$$\dot{K}(t) = \mu^{\frac{\eta}{\eta-1}}F(K,L;I^*,N) - \delta K(t) - C(t).$$

Let $\mu_N(t)$ denote the marginal value of new tasks (increasing N) in terms of the final good. Let $\mu_I(t)$ denote the marginal value of automation (increasing I) in terms of the final good. These marginal values are the social counterparts to $V_N(t)$ and $V_I(t)$ in the decentralized economy. Assuming that the planner operates in the region where $I^*(t) = I(t)$, we can write these marginal values as

$$\begin{aligned}\mu_N(t) =& (1-\eta)\mu^{\eta(1-\hat{\sigma})}\int_t^\infty e^{-\int_t^\tau (R(s)-\delta)ds}\frac{\hat{\sigma}}{1-\hat{\sigma}}Y(\tau)\left(R(\tau)^{1-\hat{\sigma}} - \gamma(n(\tau))^{\hat{\sigma}-1}w(\tau)^{1-\hat{\sigma}}\right)d\tau, \\ \mu_I(t) =& (1-\eta)\mu^{\eta(1-\hat{\sigma})}\int_t^\infty e^{-\int_t^\tau (R(s)-\delta)ds}\frac{\hat{\sigma}}{1-\hat{\sigma}}Y(\tau)\left(w(\tau)^{1-\hat{\sigma}} - R(\tau)^{1-\hat{\sigma}}\right)d\tau.\end{aligned}$$

With some abuse of notation and to maximize the parallel with the decentralized expressions for V_N and V_I, we are using $R(t)$ to denote the marginal product of capital $\mu^{\frac{\eta}{\eta-1}}F_K$ and $w(t)$ to denote the (normalized) marginal product of labor $\mu^{\frac{\eta}{\eta-1}}F_L e^{-AI^*(t)}$.

These observations show that the efficient allocation satisfies similar conditions to the decentralized economy in our main model in Section 4. The only difference is that now, the allocation of scientists is guided by $\mu_N(t)$ and $\mu_I(t)$ and satisfies:

$$S_I(t) = SG\left(\frac{\kappa_I\mu_I(t) - \kappa_N\mu_N(t)}{Y(t)}\right), \qquad S_N(t) = S\left[1 - G\left(\frac{\kappa_I\mu_I(t) - \kappa_N\mu_N(t)}{Y(t)}\right)\right],$$

so that in the efficient allocation, $n(t)$ changes endogenously according to:

$$\dot{n}(t) = \kappa_N S - (\kappa_N + \kappa_I)G\left(\frac{\kappa_I\mu_I(t) - \kappa_N\mu_N(t)}{Y(t)}\right)S.$$

One of the key insights from Proposition 6 is that the expected path for factor prices determines the incentives to automate and create new tasks. The equations for μ_N and μ_I show that a planner would also allocate scientists to developing both types of technologies following a similar principle; guided by the cost savings that each technology grants to firms. However, the fact that $\mu_N \neq V_N$ and $\mu_I \neq V_I$ shows that the decentralized allocation is not necessarily efficient. The inefficiency arises because technology monopolists do not earn the full gains that their technology generates, nor internalize how their innovations affect other existing and future technology monopolists.

We now show that labor market frictions change the planner's incentives to allocate scientists. By contrast, conditional on the wage level, such frictions do not change the market incentives to automate or create new tasks.

Without frictions, the efficient level of labor satisfies

$$\left(\mu^{\frac{\eta}{\eta-1}}F_L - c\nu'(L)\right)\mu_K \leq 0,$$

with equality if $L > 0$.

Now suppose that there is an exogenous constraint on labor that requires $L \leq L_{qs}(\omega)$. Let μ_L be the multiplier of this constraint. We have that:

$$\mu_L = \begin{cases} \left(\mu^{\frac{\eta}{\eta-1}}F_L - c\nu'(L)\right)\mu_K > 0 & \text{if } L = L_{qs}(\omega) \\ 0 & \text{if } L < L_{qs}(\omega) \end{cases}$$

Because the planner takes into account the first-order effects from changes in the employment level, the values for μ_N and μ_I change to:

$$\begin{aligned} \mu_N(t) = (1-\eta)\mu^{\eta(1-\hat{\sigma})}\int_t^{\infty} e^{-\int_t^{\tau}(R(s)-\delta)ds}\Bigg[&\frac{\hat{\sigma}}{1-\hat{\sigma}}Y(\tau)\left(R(\tau)^{1-\hat{\sigma}} - \gamma(n(\tau))^{\hat{\sigma}-1}w(\tau)^{1-\hat{\sigma}}\right) \\ &+ \left(\mu^{\frac{\eta}{\eta-1}}F_L - c\nu'(L)\right)L\frac{\widetilde{\varepsilon}_L}{\hat{\sigma}+\widetilde{\varepsilon}_L}\Lambda_N\Bigg]d\tau, \\ \mu_I(t) = (1-\eta)\mu^{\eta(1-\hat{\sigma})}\int_t^{\infty} e^{-\int_t^{\tau}(R(s)-\delta)ds}\Bigg[&\frac{\hat{\sigma}}{1-\hat{\sigma}}Y(\tau)\left(w(\tau)^{1-\hat{\sigma}} - R(\tau)^{1-\hat{\sigma}}\right) \\ &- \left(\mu^{\frac{\eta}{\eta-1}}F_L - c\nu'(L)\right)L\frac{\widetilde{\varepsilon}_L}{\hat{\sigma}+\widetilde{\varepsilon}_L}\Lambda_I\Bigg]d\tau. \end{aligned}$$

Thus, when the level of employment is below its unconstrained optimum, the planner values the introduction of new tasks more because they raise the marginal product of labor and ease the constraint on total employment. Likewise, the planner values automation less because she recognizes that by reducing employment automation has a first-order cost on workers. Importantly, the market does not recognize the first-order costs from automation or the first-order benefits from introducing new tasks. As the expressions for V_N and V_I show, only factor prices—not the extent of frictions in the labor market— determine the incentives to introduce these technologies.

When New Tasks Also Use Capital

In our baseline model, new tasks use only labor. This simplifying assumption facilitated our analysis, but is not crucial or even important for our results. Here we outline a version of the model where new tasks also use capital and show that all of our results continue to hold in this case. Suppose, in particular, that the production function for non-automated tasks is

$$y(i) = \left[\eta q(i)^{\frac{\zeta-1}{\zeta}} + (1-\eta)\left(B_\nu(\gamma(i)l(i))^\nu k(i)^{1-\nu}\right)^{\frac{\zeta-1}{\zeta}}\right]^{\frac{\zeta}{\zeta-1}}, \tag{B21}$$

where $k(i)$ is the capital used in the production of the task (jointly with labor), $\nu \in (0,1)$, and $B_\nu = \nu^{-\nu}(1-\nu)^{-(1-\nu)}$ is a constant that is re-scaled to simplify the algebra.

Automated tasks $i \leq I$ can be produced using labor or capital, and their production function takes the form

$$y(i) = \left[\eta q(i)^{\frac{\zeta-1}{\zeta}} + (1-\eta)\left(k_A(i) + B_\nu(\gamma(i)l(i))^\nu k(i)^{1-\nu}\right)^{\frac{\zeta-1}{\zeta}}\right]^{\frac{\zeta}{\zeta-1}}. \tag{B22}$$

Here $k_A(i)$ is the amount of capital used in an automated task, while $k(i)$ is the amount of capital used to produce a task with labor. Comparing these production functions to those in our baseline model (2) and (3), we readily see that the only difference is the requirement that labor has to be combined with capital in all tasks (while automated tasks continue not to use any labor). Note also that when $\nu \to 1$, we recover the model in the main text as a special case. It can be shown using a very similar analysis to that in our main model that most of the results continue to hold with minimal modifications. For example, there will exist a threshold $\widetilde{I}$ such that tasks below $I^* = \min\{I, \widetilde{I}\}$ will be produced using capital and the remaining more complex tasks will be produced using labor. Specifically, whenever $R \in \arg\min\left\{R, R^{1-\nu}\left(\frac{W}{\gamma(i)}\right)^\nu\right\}$ and $i \leq I$, the relevant task is produced using capital, and otherwise it is produced using labor. Since $\gamma(i)$ is strictly increasing, this implies that there exists a threshold $\widetilde{I}$ at which, if technologically feasible, firms would be indifferent between using capital and labor. Namely, at task $\widetilde{I}$, we have $R = W/\gamma(\widetilde{I})$, or

$$\frac{W}{R} = \gamma(\widetilde{I}).$$

This threshold represents the index up to which using capital to produce a task yields the cost-minimizing allocation of factors. However, if $\widetilde{I} > I$, firms will not be able to use capital all the way up to task $\widetilde{I}$ because of the constraint imposed by the available automation technology. For this reason, the equilibrium threshold below which tasks are produced using capital is given by

$$I^* = \min\{I, \widetilde{I}\},$$

meaning that $I^* = \widetilde{I} < I$ when it is technologically feasible to produce task $\widetilde{I}$ with capital, and $I^* = I < \widetilde{I}$ otherwise.

The demand curves for capital and labor are similar, with the only modification that the demand for capital also comes from non-automated tasks. In particular, the market-clearing conditions become:

$$K = Y(1-\nu)(1-\eta)\int_{I^*}^{N} R^{(1-\nu)(1-\zeta)-1}\left(\frac{W}{\gamma(i)}\right)^{\nu(1-\zeta)} c^u\left(R^{1-\nu}\left(\frac{W}{\gamma(i)}\right)^{\nu}\right)^{\zeta-\sigma} di \tag{B23}$$
$$+ Y(1-\eta)(I^*-N+1)c^u(R)^{\zeta-\sigma}R^{-\zeta}.$$

$$L = Y\nu(1-\eta)\int_{I^*}^{N}\frac{1}{\gamma(i)}R^{(1-\nu)(1-\zeta)}\left(\frac{W}{\gamma(i)}\right)^{\nu(1-\zeta)-1} c^u\left(R^{1-\nu}\left(\frac{W}{\gamma(i)}\right)^{\nu}\right)^{\zeta-\sigma} di. \tag{B24}$$

Following the same steps as in the text, we can then establish analogous results. This requires the more demanding Assumption 2$''$, which guarantees that the demand for factors above is homothetic:

Assumption 2$''$**:** One of the following three conditions holds:

- $\eta \to 0$;
- $\zeta \to 1$;
- or $\sigma - \zeta \to 0$.

Proposition B3 (Equilibrium in the static model when $\nu \in (0,1)$**)** *Suppose that Assumption* 1$''$ *holds. Then, for any range of tasks* $[N-1,N]$*, automation technology* $I \in (N-1,N]$*, and capital stock* K*, there exists a unique equilibrium characterized by factor prices,* W *and* R*, and threshold tasks,* $\widetilde{I}$ *and* I^**, such that: (i)* $\widetilde{I}$ *is determined by equation* (6) *and* $I^* = \min\{I,\widetilde{I}\}$*; (ii) all tasks* $i \le I^*$ *are produced using capital and all tasks* $i > I^*$ *are produced combining labor and capital; (iii) the capital and labor market-clearing conditions, equations* (B23) *and* (B24)*, are satisfied; and (iv) factor prices satisfy the ideal price index condition:*

$$(I^*-N+1)c^u(R)^{1-\sigma} + \int_{I^*}^{N} c^u\left(R^{1-\nu}\left(\frac{W}{\gamma(i)}\right)^{\nu}\right)^{1-\sigma} di = 1. \tag{B25}$$

Proof. The proof follows the same steps as Proposition 1. ∎

Comparative statics in this case are also identical to those in the baseline model (as summarized in Proposition 2) and we omit them to avoid repetition. The dynamic extension of this more general model is also very similar, and in fact, Proposition 4 applies identically, and is also omitted. One can also define $\overline{\rho}$, $\overline{n}(\rho)$ and $\widetilde{n}(\rho)$ in an analogous nway as we did in the proof of Lemma A2. To highlight the parallels, we just present the equivalent of Proposition 6.

Proposition B4 (Equilibrium with endogenous technology when $\nu \in (0,1)$**)** *Suppose that Assumptions 1$'$, 2$'$, and 4 hold. Then, there exists* $\overline{S}$ *such that, when* $S < \overline{S}$*, we have:*

1 ***(Full automation)*** *For* $\rho < \overline{\rho}$*, there is a BGP in which* $n(t) = 0$ *and all tasks are produced with capital.*

For $\rho > \overline{\rho}$, all BGPs feature $n(t) = n > \overline{n}(\rho)$. Moreover, there exist $\overline{\kappa} > \underline{\kappa} > 0$ such that:

2 ***(Unique interior BGP)*** *if $\frac{\kappa_I}{\kappa_N} > \overline{\kappa}$ there exists a unique BGP. In this BGP we have $n(t) = n \in (\overline{n}(\rho), 1)$ and $\kappa_N v_N(n) = \kappa_I v_I(n)$. If, in addition, $\theta = 0$, then the equilibrium is unique everywhere and the BGP is globally (saddle-path) stable. If $\theta > 0$, then the equilibrium is unique in the neighborhood of the BGP and is asymptotically (saddle-path) stable;*

3 ***(Multiple BGPs)*** *if $\overline{\kappa} > \frac{\kappa_I}{\kappa_N} > \underline{\kappa}$, there are multiple BGPs;*

4 ***(No automation)*** *If $\underline{\kappa} > \frac{\kappa_I}{\kappa_N}$, there exists a unique BGP. In this BGP $n(t) = 1$ and all tasks are produced with labor.*

Proof. The proof of this result closely follows that of Proposition 6, especially exploiting the fact that the behavior of profits of automation and the creation of new tasks behave identically to those in the baseline model, and thus the value functions behave identically also. ■

Microfoundations for the Quasi-Labor Supply Function

We provide various micro-foundations for the quasi-labor supply expression used in the main text, $L^s\left(\frac{W}{RK}\right)$.

Efficiency wages: Our first micro-foundation relies on an efficiency wage story. Suppose that, when a firm hires a worker to perform a task, the worker could shirk and, instead of working, use her time and effort to divert resources away from the firm.

Each firm monitors its employees, but it is only able to detect those who shirk at the flow rate q. If the worker is caught shirking, the firm does not pay wages and retains its resources. Otherwise, the worker earns her wage and a fraction of the resources that she diverted away from the firm.

In particular, assume that each firm holds a sum RK of liquid assets that the worker could divert, and that if uncaught, a worker who shirks earns a fraction $u(i)$ of this income. We assume that the sum of money that the worker may be able to divert is RK to simplify the algebra. In general, we obtain a similar quasi-supply curve for labor so long as these funds are proportional to total income $Y = R\ K + WL$.

In this formulation, $u(i)$ measures how untrustworthy worker i is, and we assume that this information is observed by firms. $u(i)$ is distributed with support $[0, \infty)$ and has a cumulative density function G. Moreover, we assume there is a mass L of workers. A worker of type $u(i)$ does not shirk if and only if:

$$W \geq (1-q)[W + u(i)RK] \rightarrow \frac{W}{RK}\frac{q}{1-q} \geq u(i).$$

Thus, when the market wage is W, firms can only afford to hire workers who are sufficiently trustworthy. The employment level is therefore given by:

$$L^s = G\left(\frac{W}{RK}\frac{q}{1-q}\right)L.$$

When $q = 1$—so that there is no monitoring problem—, we have $G\left(\frac{W}{RK}\frac{q}{1-q}\right) = 1$, and the supply of labor is fixed at L for all wages $W \geq 0$. However, when $q < 1$—so that there is a monitoring problem—, we have $L^s < L$. Even though all workers would rather work than stay unemployed, the monitoring problem implies that not all of them can be hired at the market wage. Notice that, though it is privately too costly to hire workers with $u(i) > \frac{W}{RK}\frac{q}{1-q}$, these workers strictly prefer employment to unemployment.

Alternatively, one could also have a case in which firms do not observe $u(i)$, which is private information. This also requires that firms do not learn about workers. To achieve that, we assume that workers draw a new value of $u(i)$ at each point in time.

When the marginal value of labor is W, firms are willing to hire workers so long as the market wage $\widetilde{W}$ satisfies:

$$(W - \widetilde{W})G\left(\frac{\widetilde{W}}{RK}\frac{q}{1-q}\right) - (1-q)\left(\widetilde{W} + RK\int_{\frac{\widetilde{W}}{RK}\frac{q}{1-q}}^{\infty} udG(u)\right) \geq 0.$$

This condition guarantees that the firm makes positive profits from hiring an additional worker, whose type is not known.

Competition among firms implies that the equilibrium wage at each point in time satisfies:

$$(W - \widetilde{W})G\left(\frac{\widetilde{W}}{RK}\frac{q}{1-q}\right) - (1-q)\left(\widetilde{W} + RK\int_{\frac{\widetilde{W}}{RK}\frac{q}{1-q}}^{\infty} udG(u)\right) = 0.$$

This curve yields an increasing mapping from $\frac{W}{RK}$ to $\frac{\widetilde{W}}{RK}$, which we denote by

$$\frac{\widetilde{W}}{RK} = h\left(\frac{W}{RK}\right).$$

Therefore, the effective labor supply in this economy, or the quasi-supply of labor, is given by

$$L^s = G\left(\frac{\widetilde{W}}{RK}\frac{q}{1-q}\right) = G\left(h\left(\frac{W}{RK}\right)\frac{q}{1-q}\right)L.$$

As in the previous model, even though the opportunity cost of labor is zero, the economy only manages to use a fraction of its total labor.

Minimum wages: Following Acemoglu (2003), another way in which we could obtain a quasi-labor supply curve is if there is a binding minimum wage. Suppose that the government imposes a (binding) minimum wage $\widetilde{W}$ and indexes it to the income level (or equivalently the level of consumption):

$$\widetilde{W} = \varrho \cdot (RK + WL),$$

with $\varrho > 0$. Here, $RK + WL$ represents the total income in the economy (net of intermediate goods' costs).

Suppose that the minimum wage binds. Then:

$$L = \frac{1}{\varrho}s_L,$$

which defines the quasi-labor supply in this economy as an increasing function of the labor share.

Additional References

Acemoglu, Daron (2009) *Introduction to Modern Economic Growth,* Princeton University Press.

Li, Guoying (1985) "Robust Regression," in *Exploring Data Tables, Trends, and Shapes*, pp. 281–343.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek (2017) *Integrated Public Use Microdata Series: Version 7.0 [dataset]*, Minneapolis: University of Minnesota.

Walter, Wolfgang (1998) *Ordinary Differential Equations,* Springer; Graduate Texts in Mathematics.