Online Appendix to "Uneven Growth: Automation's Impact on Income and Wealth Inequality"

Benjamin Moll, Lukasz Rachel and Pascual Restrepo

September 29, 2019

A Proofs of Propositions and Lemmas in Main Text

Proof of Lemma 1. Let $x_z(s) := w_z/r + a(s)$ denote the effective wealth of a person of age s. Equation (1) can be rewritten as:

$$\max_{\{c_z(s),x_z(s)\}_{s\geq 0}} \int_0^\infty e^{-\rho s}\frac{c_z(s)^{1-\sigma}}{1-\sigma}ds \tag{A1}$$
$$\text{s.t. } \dot{x}_z(s) = rx_z(s) - c_z(s), \text{ and } x_z(s) \geq 0$$

The Hamiltonian associated with this maximization problem is

$$H(c_z, x_z, \lambda_z) := \frac{c_z^{1-\sigma}}{1-\sigma} + \lambda(rx_z - c_z), \tag{A2}$$

where λ_z is the co-state for effective wealth.

We can write the candidate solution given in Lemma 1 as (time arguments are ignored to save on notation)

$$\dot{x}_z = \frac{r-\rho}{\sigma}x_z \qquad\qquad c_z = \left(r - \frac{r-\rho}{\sigma}\right)x_z. \tag{A3}$$

We will show that the unique solution to this system of differential equations starting from $x_z(0) = w_z/r$ solves the maximization problem in (A1).

Theorem 7.14 in Acemoglu (2009) implies that this candidate path reaches an optimum if there exists a co-state variable λ_z such that:

1. the path satisfies the restrictions $\dot{x}_z(s) = rx_z(s) - c_z(s)$, and $x_z(s) \geq 0$;
2. the following necessary conditions hold:

$$c_z^{-\sigma} = \lambda_z,$$
$$\rho\lambda_z - \dot{\lambda}_z = r\lambda_z;$$

3. the maximized Hamiltonian $M(x_z, \lambda_z) = \max_c H(c, x_z, \lambda_z)$ is concave in x_z along the candidate path;

4. the transversality condition holds. That is, for the candidate path, we have

$$\lim_{s\to\infty} e^{-\rho s} x_z \lambda_z = 0.$$

and for all other feasible paths, $\hat{x}_z$, we have

$$\lim_{s\to\infty} e^{-\rho s} \hat{x}_z \lambda_z \geq 0.$$

To prove condition 1, note that starting from any $x_z(0) \geq 0$, we will have $x_z \geq 0$. Moreover, for any path satisfying equations (A3) the flow budget constraint holds:

$$\begin{aligned} r x_z - c_z =& r x_z - \left(r - \frac{r-\rho}{\sigma}\right) x_z \\ =& \frac{r-\rho}{\sigma} x_z \\ =& \dot{x}_z. \end{aligned}$$

To prove condition 2, define $\lambda_z := (r - (r-\rho)/\sigma)^{-\sigma} x_z^{-\sigma} > 0$ (here we used the condition $r > (r-\rho)/\sigma$). By construction, $c_z^{-\sigma} = \lambda_z$. Moreover:

$$\begin{aligned} \rho\lambda_z - \dot{\lambda}_z =& \rho \left(r - \frac{r-\rho}{\sigma}\right)^{-\sigma} x_z^{-\sigma} + \left(r - \frac{r-\rho}{\sigma}\right)^{-\sigma} \sigma x_z(s)^{-\sigma-1} \dot{x}_z \\ =& \left(\rho + \sigma \frac{\dot{x}_z}{x_z}\right) \left(r - \frac{r-\rho}{\sigma}\right)^{-\sigma} x_z(s)^{-\sigma} \\ =& \left(\rho + \sigma \frac{\dot{x}_z}{x_z}\right) \lambda_z \\ =& \left(\rho + \sigma \frac{r-\rho}{\sigma}\right) \lambda_z \\ =& r\lambda_z. \end{aligned}$$

To prove condition 3, note that

$$\max_c H(c, x_z, \lambda_z) = \frac{\lambda_z^{\frac{\sigma-1}{\sigma}}}{1-\sigma} + \lambda_z (r x_z - \lambda_z^{-\frac{1}{\sigma}}),$$

which is concave (linear) in x_z.

To prove the first part of condition 4, note that along the candidate path, x_z grows at a rate $\frac{r-\rho}{\sigma}$, and λ_z at a rate $\rho - r$. It follows that the first part of the transversality condition holds if

$$-\rho + \frac{r-\rho}{\sigma} + \rho - r < 0,$$

which is equivalent to the condition $r > (r-\rho)/\sigma$.

The second part of the transversality condition follows from the fact that, along any feasible path, we have $\hat{x}_z \geq 0$.

It follows that the candidate paths given in Lemma 1 provide optimal paths for consumption and asset accumulation in a steady state. ■

Proof of Lemma 2. Denote by $p_z(u)$ the price of task $\mathcal{Y}_z(u)$, and by p_z the price of sector z output Y_z. Cost minimization in the production of sector z output implies that the quantity of task u used is given by

$$\mathcal{Y}_z(u) = \frac{p_z Y_z}{p_z(u)}$$

Assumption 1 implies that all tasks $u \in [0, \alpha_z]$ are produced with capital. It follows that for those tasks, $p_z(u) = R$ and the quantity of capital required to produce $\mathcal{Y}_z(u)$ is $p_z Y_z / R$. It follows that the total amount of capital used in sector z is:

$$K_z = \frac{\alpha_z p_z Y_z}{R} \tag{A4}$$

Assumption 1 implies that all tasks $u \in (\alpha_z, 1]$ are produced with labor. It follows that for those tasks, $p_z(u) = \frac{w_z}{\psi_z}$ and the quantity of labor required to produce $\mathcal{Y}_z(u)$ is $p_z Y_z / w_z$. It follows that the total amount of labor of skill z used in sector z is:

$$\ell_z = \frac{(1-\alpha_z) p_z Y_z}{w_z} \tag{A5}$$

With perfect competition, the price of sector z output equals the marginal cost of production. Because tasks are combined via a Cobb-Douglas aggregator, the price is given by the dual $\ln(p_z) = \int_0^1 \ln(p_z(u)) du$. It follows that

$$p_z = R^{\alpha_z} \left(\frac{w_z}{\psi_z}\right)^{1-\alpha_z}. \tag{A6}$$

Combining the formula for p_z in (A6) with capital and labor demand conditions (A4) and (A5) gives the production of sector z as a function of the total capital and labor used in this sector, K_z and ℓ_z:

$$Y_z = \left(\frac{K_z}{\alpha_z}\right)^{\alpha_z} \left(\frac{\psi_z \ell_z}{1-\alpha_z}\right)^{1-\alpha_z} \tag{A7}$$

We now turn to aggregate output. We normalize the price of the final good to 1, so that the demand for sector z output satisfies $p_z Y_z = \gamma_z Y$.

Using equations (A4), we can compute the demand for capital in sector z as

$$K_z = \frac{\alpha_z p_z Y_z}{R} = \frac{\alpha_z \gamma_z Y}{R}.$$

Adding this formula across sectors, it follows that the total amount of capital used in the economy is

$$K = \alpha \frac{Y}{R}, \tag{A8}$$

where recall that $\alpha := \sum_z \alpha_z \gamma_z$. The share of capital allocated to sector z is therefore equal to

$$K_z = K \frac{\alpha_z \gamma_z}{\alpha}. \tag{A9}$$

Substituting this formula into (A7) we get:

$$Y_z = \left(K \frac{\gamma_z}{\alpha} \right)^{\alpha_z} \left(\frac{\psi_z L_z}{1 - \alpha_z} \right)^{1-\alpha_z}. \tag{A10}$$

Substituting sectoral outputs into the aggregate production function we obtain the formula in equation (3), with

$$\mathcal{A} := A \alpha^{-\alpha} \prod_z (\gamma_z (1 - \alpha_z))^{-\gamma_z (1-\alpha_z)} \prod_z \gamma_z^{\gamma_z}$$

The formula for wages follows from equation (A5):

$$w_z = \frac{(1 - \alpha_z) p_z Y_z}{\ell_z} = \frac{(1 - \alpha_z) \gamma_z Y}{\ell_z}.$$

The formula for the rental rate of capital follows from equation (A8). ■

Proof of the expression for productivity gains from automation. The formula in equation (3) can be written as

$$Y = A \cdot \prod_z \gamma_z^{\gamma_z} \cdot \left(\frac{K}{\alpha} \right)^{\alpha} \cdot \prod_z \left(\frac{\psi_z \ell_z}{\gamma_z (1 - \alpha_z)} \right)^{\gamma_z (1-\alpha_z)}.$$

It follows that

$$\begin{aligned}\frac{d\ln Y}{d\alpha_z}|_{K,\ell_z} =&\gamma_z \ln K - \gamma_z \ln\alpha - \gamma_z \ln(\psi_z\ell_z) + \gamma_z \ln(\gamma_z(1-\alpha_z)) \\ =&\gamma_z \ln\left(\frac{K}{\alpha}\right) - \ln\left(\frac{\psi_z\ell_z}{\gamma_z(1-\alpha_z)}\right) \\ =&\gamma_z \ln\left(\frac{Y}{R}\right) - \ln\left(\psi_z\frac{Y}{w_z}\right) \\ =&\gamma_z \ln\left(\frac{w_z}{\psi_z R}\right) > 0.\end{aligned}$$

The third row uses the factor price formulas in equation (4). ■

Lemma A1 (Lemma ensuring adoption of automation technologies) *Suppose that for all z, the following inequality holds:*

$$(\rho + p\sigma + \delta)^{-\frac{1}{1-\alpha}}\mathcal{A}^{\frac{1}{1-\alpha}}\alpha^{\frac{\alpha}{1-\alpha}} > \frac{1}{(1-\alpha_z)\gamma_z}\frac{\ell_z\psi_z}{\prod_v(\ell_v\psi_v)^{\frac{\gamma_v(1-\alpha_v)}{1-\alpha}}}. \tag{A11}$$

The equilibrium will involve the adoption of all available automation technologies.

The above inequality holds for values of A above a threshold $\bar{A}$.

Proof. We assume that all automation technologies are adopted and verify that in equilibrium, the condition above ensures that $w_z^* > \psi_z R^*$.

In order to get finite accumulation of wealth, the individual problem requires $r^* < \rho + p\sigma$. using the formula for factor prices in equation (4), we obtain:

$$(K/Y)^* > \frac{\alpha}{\rho + p\sigma + \delta}.$$

Using the fact that $Y = \mathcal{A}^{\frac{1}{1-\alpha}}(K/Y)^{\frac{\alpha}{1-\alpha}}\prod_v(\psi_v\ell_v)^{\frac{\gamma_z(1-\alpha_z)}{1-\alpha}}$, and $w_z = (1-\alpha_z)\frac{\gamma_z}{\ell_z}Y$, this inequality implies

$$w_z^* > \frac{(1-\alpha_z)\gamma_z}{\ell_z}\mathcal{A}^{\frac{1}{1-\alpha}}\alpha^{\frac{\alpha}{1-\alpha}}\prod_z(\psi_z\ell_z)^{\frac{\gamma_z(1-\alpha_z)}{1-\alpha}} \times (\rho + p\sigma + \delta)^{-\frac{\alpha}{1-\alpha}}.$$

Because $\psi_z(\rho + p\sigma + \delta) > \psi_z R^*$, It follows that a sufficient condition to ensure $w_z^* > \psi_z R^*$ is that

$$\frac{(1-\alpha_z)\gamma_z}{\ell_z}\mathcal{A}^{\frac{1}{1-\alpha}}\alpha^{\frac{\alpha}{1-\alpha}}\prod_z(\psi_z\ell_z)^{\frac{\gamma_z(1-\alpha_z)}{1-\alpha}} \times (\rho + p\sigma + \delta)^{-\frac{\alpha}{1-\alpha}} > \psi_z(\rho + p\sigma + \delta).$$

This inequality is equivalent to (A11). To conclude the proof of the lemma, note that (A11) holds for large values of A. ■

Proof of Proposition 1.

The main text presents the derivations of the supply curve (equation (8)) and the demand curve (equation (9)).

The supply curve $(K/\bar{w})^s$ increases from zero to infinity as r increases from ρ to $\rho + p\sigma$. For $r < \rho$ individuals supply no capital. For $r > \rho + p\sigma$, individuals amaze a divergent amount of capital.

The demand curve $(K/\bar{w})^s$ decreases from $(\alpha/(1-\alpha))/(\rho+\delta) > 0$ to $(\alpha/(1-\alpha))/(\rho + p\sigma + \delta) > 0$ as r increases from ρ to $\rho + p\sigma$.

These observations imply that equation (6) has a unique solution r^* and that this unique solution lies in $(\rho, \rho + p\sigma)$. In fact, r^* can be computed analytically as

$$r^* = \frac{-((1-\alpha)\delta - \rho - \alpha p\sigma) + \sqrt{((1-\alpha)\delta - \rho - \alpha p\sigma)^2 + 4(1-\alpha)\rho\delta}}{2} \tag{A12}$$

The equilibrium return r^* determines the remaining aggregates as follows. First, the capital-output ratio is given by

$$(K/Y)^* = \frac{\alpha}{r^* + \delta}.$$

The output level is given by

$$Y^* = \mathcal{A}^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{r^* + \delta}\right)^{\frac{\alpha}{1-\alpha}} \prod_z (\ell_z \psi_z)^{\frac{\gamma_z(1-\alpha_z)}{1-\alpha}}.$$

These two equations combined imply

$$K^* = \mathcal{A}^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{r^* + \delta}\right)^{\frac{1}{1-\alpha}} \prod_z (\ell_z \psi_z)^{\frac{\gamma_z(1-\alpha_z)}{1-\alpha}}. \tag{A13}$$

Turning to wages, equation (4) implies:

$$w_z^* = (1-\alpha_z)\frac{\gamma_z}{\ell_z} \mathcal{A}^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{r^* + \delta}\right)^{\frac{\alpha}{1-\alpha}} \prod_z (\ell_z \psi_z)^{\frac{\gamma_z(1-\alpha_z)}{1-\alpha}}.$$

Finally, equation (7) can be derived from the household side. As explained in the main text, in steady state we must have

$$0 = \frac{r^* - \rho}{\sigma}\left(K^* + \frac{\bar{w}^*}{r^*}\right) - pK^*.$$

This expression can be rearranged as

$$r^* = \rho + p\sigma \frac{r^* K^*}{r^* K^* + \bar{w}^*} = \rho + p\sigma\alpha^*_{net}.$$

Note that the condition $r^* > (r^* - \rho)/\sigma$, which is needed to ensure the individuals' policy functions are an optimum is equivalent to $\rho + p\alpha^*_{net}(\sigma - 1) > 0$. ■

Proof of Proposition 2. We can rearrange (6) as:

$$\frac{\left(1 - \frac{\rho}{r^*}\right)(r^* + \delta)}{p\sigma + \rho - r^*} = \frac{\alpha}{1-\alpha}.$$

The right hand side of this equation is increasing in r^*, and the left is increasing in α. It follows that r^* is increasing in α.

The individual accumulation rate is given by $(r^* - \rho)/\sigma$, and so it also increases with α.

The net capital share satisfies the identity in equation (7), and so it also increases with α.

Denote the capital-output ratio by κ^*. Equation (7) can be written as

$$r^* = \rho + p\sigma \frac{\alpha - \delta\kappa^*}{1 - \delta\kappa^*}.$$

Plugging this expression into equation (4), we obtain

$$\kappa^* \left(\rho + \delta + p\sigma \frac{\alpha - \delta\kappa^*}{1 - \delta\kappa^*}\right) = \alpha.$$

Isolating α from this equation yields

$$\kappa^* \left(\rho + \delta + \frac{p\rho\sigma\kappa^*}{1 - \delta\kappa - p\sigma\kappa^*}\right) = \alpha. \tag{A14}$$

This equation defines κ^* implicitly as a an increasing function of α, since the left-hand side increases in κ^*.

Finally, turning to output, from equation (3) we have

$$d\ln Y^* = \frac{1}{1-\alpha} d\ln \mathrm{TFP}_\alpha + \frac{\alpha}{1-\alpha} d\ln(K/Y)^*.$$

Because $d\ln \mathrm{TFP}_\alpha > 0$ and $d\ln(K/Y)^* > 0$ following any increase in the α_z's, we have that automation always increase output. ■

Proof of Proposition 3. Equation (4) implies that relative wages satisfy

$$\frac{w_z}{w_v} = \frac{(1-\alpha_z)}{(1-\alpha_v)}\frac{\gamma_z}{\gamma_v}\frac{\ell_v}{\ell_z}.$$

It follows that an increase in α_z reduces w_z/w_v for $v \neq z$.

Turning to the wage bill, note that we can add up equation (4) for all z to obtain:

$$\bar{w} = (1-\alpha)Y.$$

It follows that

$$d\ln \bar{w} = -\frac{1}{1-\alpha}\sum \gamma_z d\alpha_z + \frac{1}{1-\alpha} d\ln \mathrm{TFP}_\alpha + \frac{\alpha}{1-\alpha} d\ln(K/Y)^*.$$

We now show that the terms $d\ln \mathrm{TFP}_\alpha$ and $d\ln(K/Y)^*$ are both decreasing in p and converge to zero as p increases. Because the term $-(1/(1-\alpha_z))\sum \gamma_z d\alpha_z$ is negative, this establishes the existence of the threshold $\bar{p}$.

We first analyze the term $d\ln \mathrm{TFP}_\alpha$. This is given by

$$\begin{aligned} d\ln \mathrm{TFP}_\alpha = &\sum_z \gamma_z \ln\left(\frac{w_z^*}{\psi_z R^*}\right) \\ &\sum_z \gamma_z \ln\left(\frac{K^*}{\alpha}\right) - \gamma_z \ln\left(\frac{\psi_z \ell_z}{\gamma_z(1-\alpha_z)}\right), \end{aligned}$$

where we used the formulas for factor prices in equation (4). It is enough to show that K^* is decreasing in p and that K^* converges to zero as p increases. Because K^* is given by (A13), it is enough to show that r^* is increasing in p and that r^* converges to infinity as p increases.

The fact that r^* increases in p follows from equation (6). An increase in p contracts the supply of capital, which results on a higher r^*. Moreover, equation (A12) shows that $r^* \to \infty$ as $p \to \infty$. Note that the formal limit of $d\ln \mathrm{TFP}_\alpha$ as $p \to \infty$ is zero, since as K^* declines, we eventually reach a point where Assumption 1 starts failing and increases in α_z do not affect productivity.

We now turn to the term $d\ln(K/Y)^*$. As above, denote the capital-output ratio by κ. Applying the implicit function theorem to equation (A14), we get

$$d\ln(K/Y)^* = \frac{\alpha\rho}{\alpha\rho + p\sigma {\alpha_{net}^*}^2}\frac{d\alpha}{\alpha}.$$

As p increases we have that α^*_{net} increases (recall that $\alpha^*_{net} = (\alpha - \delta(K/Y)^*)/(1-\delta(K/Y)^*)$, which decreases in $(K/Y)^*$ and hence increases in p) and the elasticity of $(K/Y)^*$ with respect to α converges to zero.

The above argument shows that there exists some $\bar{p}$ such that, for $p > \bar{p}$, $d\ln\bar{w} < 0$. To conclude the proof, we show that $\bar{p} > 0$. This follows from the fact that, for $p = 0$, $d\ln\bar{w} > 0$.

To show this, note that for $p = 0$ we get

$$d\ln(K/Y)^* = \frac{d\alpha}{\alpha} = \frac{1}{\alpha}\sum_z \gamma_z d\alpha_z,$$

and therefore

$$\begin{aligned} d\ln\bar{w} =& -\frac{1}{1-\alpha}\sum \gamma_z d\alpha_z + \frac{1}{1-\alpha} d\ln \mathrm{TFP}_\alpha + \frac{\alpha}{1-\alpha}\frac{1}{\alpha}\sum_z \gamma_z d\alpha_z \\ =& \frac{1}{1-\alpha} d\ln\mathrm{TFP}_\alpha > 0. \end{aligned}$$

■

Proof of Proposition 4. Below we derive the effective wealth distribution, the wealth distribution, and the income distribution. To save on notation, we do not include asterisks when denoting steady state objects.

Effective wealth distribution: Denote the stationary density of effective wealth conditional on a given skill type z by $g_z(x)$. g_z satisfies the Kolmogorov Forward Equation (KFE)

$$0 = -\left(\frac{r-\rho}{\sigma} x g_z(x)\right)' - p g_z(x)$$

on $(w_z/r, \infty)$. We guess and verify that g is Pareto, i.e. $g_z(x) = c\zeta x^{-\zeta-1}$ for some constants c and ζ. Substituting in the guess

$$\begin{aligned} 0 &= \zeta\frac{r-\rho}{\sigma} c\zeta x^{-\zeta-1} - pc\zeta x^{-\zeta-1} \\ 0 &= \zeta\frac{r-\rho}{\sigma} - p \\ \frac{1}{\zeta} &= \frac{r-\rho}{p\sigma} = \alpha_{net} \end{aligned}$$

Since $g_z(x) = c\zeta x^{-\zeta-1}$ must integrate to 1 on $(w_z/r, \infty)$, we must have $c = (w_z/r)^{-\zeta}$. Hence this is a Pareto distribution with tail parameter $\zeta = \frac{1}{\alpha_{net}}$ and scale parameter $x_z(0) = w_z/r$.

Because the distribution of effective wealth is Pareto, the conditional counter-CDF for effective wealth of each skill type z is of the form:

$$\Pr(\text{effective wealth} \geq x|z) = \left(\frac{x}{w_z/r}\right)^{-\frac{1}{\alpha_{net}}}, \quad x \geq w_z/r. \tag{A15}$$

Wealth distribution: We now derive the counter-CDF for wealth. Recall that effective wealth x is $x := a + w_z/r$. Therefore

$$\Pr(\text{wealth} \geq a|z) = \Pr(\text{effective wealth} \geq a + w_z/r|z) = \left(\frac{a + w_z/r}{w_z/r}\right)^{-\frac{1}{\alpha_{net}}}, \quad a \geq 0$$

To find the unconditional distribution, we add across the different skill-types, which yields

$$\Pr(\text{assets} \geq a) = \sum_z \ell_z \left(\frac{a + w_z/r}{w_z/r}\right)^{-\frac{1}{\alpha_{net}}}.$$

Income Distribution: We now derive the counter-CDF for income. The income of a person with effective wealth x is rx. Therefore

$$\Pr(\text{income} \geq y|z) = \Pr(\text{effective wealth} \geq y/r|z) = \left(\frac{y/r}{w_z/r}\right)^{-\frac{1}{\alpha_{net}}}, \quad y \geq w_z.$$

To find the unconditional distribution, we add across the different skill-types, which yields

$$\Pr(\text{income} \geq y) = \sum_z \left(\frac{\max\{y, w_z\}}{w_z}\right)^{-\frac{1}{\alpha_{net}}}.$$

Finally, when $\delta = 0$, we have $\alpha_{net} = \alpha$ and $\frac{1}{\zeta} = \alpha$. When $\delta > 0$, Proposition 2, implies that $\frac{1}{\zeta}$ is increasing in α. ∎

Proof of Proposition 5. We start by deriving the probability that individuals with skill z are among the top q income earners. To save on notation, we do not include asterisks when denoting steady state objects.

Let $y(q)$ denote the income of the qth higher earner. That is:

$$\Pr(\text{income} \geq y(q)) = q.$$

By definition, for $q < \bar{q}$ we have $y(q) > \max_z w_z$. We can thus compute $y(q)$ explicitly as

$$q = \Pr(\text{income} \geq y(q)) = \sum_z \ell_z \left(\frac{y(q)}{w_z}\right)^{-1/\alpha_{net}},$$

which implies

$$y(q) = q^{-\alpha_{net}} \left(\sum_z \ell_z w_z^{1/\alpha_{net}} \right)^{\alpha_{net}}. \tag{A16}$$

An application of Bayes' rule implies

$$\Pr(\text{skill} = z|\text{top } q) = \frac{\ell_z \Pr(\text{income} \geq y(q)|z)}{\sum_v \ell_v \Pr(\text{income} \geq y(q)|v)} = \frac{\ell_z y(q)^{-1/\alpha_{net}} w_z^{1/\alpha_{net}}}{\sum_v \ell_v y(q)^{-1/\alpha_{net}} w_v^{1/\alpha_{net}}} = \frac{\ell_z w_z^{1/\alpha_{net}}}{\sum_v \ell_v w_v^{1/\alpha_{net}}}.$$

We now turn to the share of labor income at the top of the income distribution. The expected labor income for individuals at the top q is given by

$$\mathbb{E}[\text{labor income}|\text{top } q] = \sum_z w_z \Pr(\text{skill} = z|\text{top } q) = \sum_z \frac{\ell_z w_z^{1+1/\alpha_{net}}}{\sum_v \ell_v w_v^{1/\alpha_{net}}} = \frac{\sum_z \ell_z w_z^{1+1/\alpha_{net}}}{\sum_z \ell_z w_z^{1/\alpha_{net}}}.$$

The expected income for individuals at the top q is given by

$$\mathbb{E}[\text{income}|\text{top } q] = \sum_z \mathbb{E}[\text{income}|\text{income} \geq y(q)] \Pr(\text{skill} = z|\text{top } q) = \frac{y(q)}{1 - \alpha_{net}}.$$

Here, we used the fact that $\mathbb{E}[\text{income}|\text{income} \geq y(q)] = \frac{y(q)}{1-\alpha_{net}}$, a well-known property of Pareto distributions.

It follows that

$$\frac{\mathbb{E}[\text{labor income}|\text{top } q]}{\mathbb{E}[\text{income}|\text{top } q]} = \frac{1 - \alpha_{net}}{y(q)} \frac{\sum_z \ell_z w_z^{1+1/\alpha_{net}}}{\sum_z \ell_z w_z^{1/\alpha_{net}}} = (1 - \alpha_{net}) q^{\alpha_{net}} \frac{\sum_z \ell_z w_z^{1+1/\alpha_{net}}}{\left(\sum_z \ell_z w_z^{1/\alpha_{net}}\right)^{1+\alpha_{net}}}$$

Finally, we compute the share of national income earned by the top q. For $q \leq \bar{q}$, the top q earn an income

$$T(q) = q\mathbb{E}[\text{income}|\text{top } q] = \frac{1}{1 - \alpha_{net}} q^{1-\alpha_{net}} \left(\sum_z \ell_z w_z^{1/\alpha_{net}} \right)^{\alpha_{net}}.$$

It follows that the top q earn a share of national income equal to:

$$S(q) = \frac{S(q)}{S(\bar{q})} S(\bar{q}) = \frac{T(q)}{T(\bar{q})} S(\bar{q}) = \frac{q^{1-\alpha_{net}}}{\bar{q}^{1-\alpha_{net}}} S(\bar{q}).$$

The result in the proposition follows by letting $\Lambda = \frac{1}{\bar{q}^{1-\alpha_{net}}} S(\bar{q})$. ■

Proof of Proposition 6. Let $q < \bar{q}$, with $\bar{q}$ defined in Proposition 5.

From Proposition 5, it follows that $S(q) = \Lambda q^{1-\alpha_{net}}$. Moreover, because the skill composition of workers among the top q is constant, $\tilde{S}^\ell(q) = \Lambda^\ell q$, where $\Lambda^\ell = \frac{1}{\bar{q}}\tilde{S}^\ell(\bar{q})$. It follows that

$$S(q) = \Lambda q^{1-\alpha_{net}}, \qquad \tilde{S}^\ell(q) = \Lambda^\ell q, \qquad \tilde{S}^k(q) = \frac{1}{\alpha_{net}}\left(\Lambda q^{1-\alpha_{net}} - (1-\alpha_{net})\Lambda^\ell q\right).$$

The compositional effect is then given by

$$\text{Compositional effect at } q = \frac{1}{\alpha^*_{net}}\left(\Lambda q^{1-\alpha^*_{net}} - \Lambda^\ell q\right) d\alpha^*_{net} > 0,$$

whereas the overall change in the share of income held by the top q percent is

$$\text{Total change at } q = \Lambda \ln(1/q) \times q^{1-\alpha^*_{net}} d\alpha^*_{net} > 0.$$

The share of the total change in $S(q)$ explained by the compositional effect is then given by

$$\frac{\Lambda - \Lambda^\ell q^{\alpha^*_{net}}}{\Lambda \alpha^*_{net} \ln(1/q)},$$

which converges to zero as $q \to 0$. ■

B Outline and Proofs for the Full Model with Transitional Dynamics

This section describes the full model in a non-stationary environment. Because firms rent capital from households, the production structure remains unchanged. We therefore focus on the individual problem in a non-stationary environment. We also present a explicit microfoundation behind the assumption that individuals consume their wealth upon death. The section concludes with the proof of Proposition 7.

Savings Problem in a Non-stationary Environment

During their lives, with Poisson rate p, individuals are told they will die in T periods. We work with the limit as $T \to 0$, i.e. you only have one more millisecond to live. Finally, to capture the possibility of imperfect dynasties, we assume that individuals do not derive any utility from bequests.

Let $V^T_z(t; a)$ denote the value of lifetime consumption for a person who enters the dying state at time t with assets a. The following lemma characterizes the behavior of this value function and the consumption of individuals in the dying state.

Lemma A2 (Characterization of the dying state) *For $T \to 0$, the value function $V_z^T(a,t)$ converges to zero*

$$\lim_{T \to 0} V_z^T(a,t) \to 0 \quad \forall a.$$

In this limit case, the flow consumption of the dying is $pA(t)$, where $A(t)$ are total asset holdings in the economy.

Before proceeding with the Lemma's proof, we briefly explain the intuition for the result that $V_z^T(a,t) \to 0$ as $T \to 0$. As already stated in footnote 10, individuals enjoy an infinite end-of-life consumption flow over an infinitesimal time interval $T \to 0$. But because utility is strictly concave, the value of such end-of-life consumption converges to zero as $T \to 0$. In terms of the notation in the proof below, we have

$$V_z^T(t;a) \approx \int_0^T u(a/T)dt = T\frac{(a/T)^{1-\sigma}}{1-\sigma} = T^\sigma \frac{a^{1-\sigma}}{1-\sigma} \to 0 \quad \text{as} \quad T \to 0.$$

Proof. People with wealth a that learn that they will die in T periods solve the following problem:

$$V_z^T(t;a) = \max_{\{c_z^T(t+\tau;a), a_z^T(t+\tau;a)\}} \int_0^T e^{-\varrho\tau} \frac{c_z^T(t+\tau;a)^{1-\sigma}}{1-\sigma} d\tau \tag{A17}$$
$$\text{s.t.} \quad \dot{a}_z^T(t+\tau;a) = w_z(t) + r(t)a_z^T(t+\tau;a) - c_z^T(t+\tau;a), \quad a_z^T(T;a) \geq 0, \quad a_z^T(0;a) = a.$$

The Euler equation for this problem implies that consumption satisfies

$$c_z^T(t+\tau;a) = c_z^T(t;a) \times \exp\left(\int_0^\tau \frac{1}{\sigma}(r(t+\chi) - \varrho)d\chi\right). \tag{A18}$$

Because $a_z^T(T) > 0$, any optimal consumption path must involve $a_z^T(T) = 0$. Therefore, the consumption path satisfies the budget constraint

$$\int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) c_z^T(t+\tau;a)d\tau = a + \int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) w_z(t+\tau)d\tau$$

Plugging equation (A18) into the budget constraint, we obtain the following expression for $c_z^T(t;a)$:

$$c_z^T(t;a) = \frac{a + \int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) w_z(t+\tau)d\tau}{\int_0^T \exp\left(-\int_0^\tau (r(t+\chi)(1-1/\sigma) + \varrho/\sigma)d\chi\right) d\tau}.$$

Thus, the optimal consumption path satisfies

$$c_z^T(t+\tau;a) = \frac{a + \int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) w_z(t+\tau)d\tau}{\int_0^T \exp\left(-\int_0^\tau (r(t+\chi)(1-1/\sigma)+\varrho/\sigma)d\chi\right)d\tau} \times \exp\left(\int_0^\tau \frac{1}{\sigma}(r(t+\chi)-\varrho)d\chi\right). \tag{A19}$$

Plugging equation (A19) into (A17), we obtain

$$\begin{aligned} V_z^T(t;a) =& \frac{1}{1-\sigma}\left(a + \int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) w_z(t+\tau)d\tau\right)^{1-\sigma} \times \\ & \left(\int_0^T \exp\left(-\int_0^\tau (r(t+\chi)(1-1/\sigma)+\varrho/\sigma)d\chi\right)d\tau\right)^{\sigma} - \frac{1}{1-\sigma}\int_0^T e^{-\varrho\tau}d\tau. \end{aligned}$$

As $T \to 0$, we have that $a + \int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) w_z^T(t+\tau)d\tau$ converges to a and $\int_0^T \exp\left(-\int_0^\tau (r(t+\chi)(1-1/\sigma)+\varrho/\sigma)d\chi\right)d\tau$ and $\int_0^T e^{-\varrho\tau}d\tau$ converge to zero. It follows that $V_z^T(t;a)$ converges to zero, as claimed in the lemma.

We now turn to computing the flow consumption of individuals in the dying state. For $\tau \in [0,T]$, let $h_z(a,t-\tau)$ denote the mass of individuals of skill z who entered the dying state with assets a at time $t-\tau$. We have that the flow consumption of the dying equals

$$\begin{aligned} C_D^T(t) =& \sum_z \int_a \int_0^T c_z^T(t+\tau;a)h_z(a,t-\tau)d\tau da \\ =& \sum_z \int_a \int_0^T \frac{a + \int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) w_z(t+\tau)d\tau}{\int_0^T \exp\left(-\int_0^\tau (r(t+\chi)(1-1/\sigma)+\varrho/\sigma)d\chi\right)d\tau} \times \\ & \qquad \exp\left(\int_0^\tau \frac{1}{\sigma}(r(t+\chi)-\varrho)d\chi\right) h_z(a,t-\tau)d\tau da \\ =& \sum_z \int_a \left(a + \int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) w_z(t+\tau)d\tau\right) \times \\ & \qquad \left(\frac{\int_0^T \exp\left(\int_0^\tau \frac{1}{\sigma}(r(t+\chi)-\varrho)d\chi\right) h_z(a,t-\tau)d\tau}{\int_0^T \exp\left(-\int_0^\tau (r(t+\chi)(1-1/\sigma)+\varrho/\sigma)d\chi\right)d\tau}\right) da. \end{aligned}$$

To compute the limit of $C_D^T(t)$ as $t \to 0$, note that

$$\lim_{T\to 0} a + \int_0^T \exp\left(-\int_0^\tau r(t+\chi)d\chi\right) w_z(t+\tau)d\tau = a,$$

and, by an application of L'Hôpital's rule

$$\lim_{T\to 0}\frac{\int_0^T \exp\left(\int_0^\tau \frac{1}{\sigma}(r(t+\chi)-\varrho)d\chi\right)d\tau}{\int_0^T \exp\left(-\int_0^\tau (r(t+\chi)(1-1/\sigma)+\varrho/\sigma)d\chi\right)d\tau} = h_z(a,t).$$

It follows that

$$\lim_{T\to 0} C_D^T(t) = \sum_z \int_a a h_z(a,t)da = pA(t).$$

The last step follows from the fact that the probability of entering the dying state is p, independently of assets and skills. ■

In what follows, we work with the limit $T \to 0$. Note that the above argument applies for individuals who reach the dying state with non-negative assets (the above problem when $T \to 0$ is only well defined when $a > 0$). We assume that individuals who reach this state with negative assets are forced to engage in negative consumption in order to repay their debt and obtain zero utility from T onwards.

Lemma A2 and the remarks above imply that consumption and saving decisions of individuals who are not in the dying state, solve the following maximization problem, which generalizes (1) in the main text:

$$\max_{\{c_z(t,b),a_z(t,b)\}_{t\geq b}} \int_0^\infty e^{-(\varrho+p)t}\frac{c_z(t,b)^{1-\sigma}}{1-\sigma}ds \tag{A20}$$
$$\text{s.t. } \dot{a}_z(t,b) = w_z(t) + r(t)a_z(t,b) - c_z(t,b), \text{ and } a_z(t,b) \geq \int_t^\infty e^{-\int_t^\tau r(\chi)d\chi} w_z(\tau)d\tau$$

Here, $c_z(t,b)$ and $a_z(t,b)$ denote the consumption and assets of a person of skill type z from the cohort born at time b at time t (so that they age is $t-b$). Unlike in the main text, it is convenient to keep track of time and birth cohort, rather than age and time. The extra discounting $\varrho+p$ captures the possibility of entering the dying state where individuals derives no utility.

To characterize the solution to this problem, we generalize the definition of effective wealth to

$$x_z(t,b) := a_z(t,b) + h_z(t),$$

where $h_z(t)$ denotes the human wealth of individuals with skill z, given by

$$h_z(t) := \int_t^\infty e^{-\int_t^\tau r(\tau)d\tau} w_z(s)ds.$$

Lemma A3 (Households policy functions outside the steady state) *Suppose that $r(t)$ converges to r^* and $r^* > (r^* - \rho)/\sigma$. The unique solution to the individual problem in equation (A20) is given by policy functions that are linear in effective wealth*

$$\begin{aligned} \dot{x}_z(t,b) &= (r(t) - \mu(t))x_z(t,b), \\ c_z(t,b) &= \mu(t)x_z(t,b), \end{aligned} \tag{A21}$$

for $t \geq b$, with $x_z(b,b) = h_z(b)$.

Here $\mu(t)$ denotes the marginal propensity to consume out of wealth, and satisfies the differential equation:

$$\frac{\dot{\mu}(t)}{\mu(t)} = \mu(t) - r(t) + \frac{1}{\sigma}(r(t) - \rho) \tag{A22}$$

Proof. The maximization problem can be rewritten using effective wealth as

$$\max_{\{c_z(t,b),x_z(t,b)\}_{t\geq b}} \int_0^\infty e^{-(\varrho+p)s} \frac{c_z(t,b)^{1-\sigma}}{1-\sigma} dt \tag{A23}$$
$$\text{s.t. } \dot{x}_z(t,b) = r(t)x_z(t,b) - c_z(t,b), \text{ and } x_z(t,b) \geq 0$$

The Hamiltonian associated with this maximization problem is

$$H(c_z, x_z, \lambda_z) := \frac{c_z^{1-\sigma}}{1-\sigma} + \lambda(rx_z - c_z), \tag{A24}$$

where λ_z is the co-state for effective wealth.

We show that the unique solution to equation (A21) starting from $x_z(b,b) = \int_b^\infty e^{-\int_t^\tau r(\chi)d\chi} w_z(\tau)d\tau$ solves the maximization problem in (A1).

Theorem 7.14 in Acemoglu (2009) implies that this candidate path reaches an optimum if there exists a co-state variable $\lambda_z(t,b)$ such that:

1. the path satisfies the restrictions $\dot{x}_z(t,b) = r(t)x_z(t,b) - c_z(t,b)$, and $x_z(t,b) \geq 0$;
2. the following necessary conditions hold:

$$\begin{aligned} c_z(t,b)^{-\sigma} &= \lambda_z(t,b), \\ \rho\lambda_z(t,b) - \dot{\lambda}_z(t,b) &= r\lambda_z(t,b); \end{aligned}$$

3. the maximized Hamiltonian $M(x_z, \lambda_z) = \max_c H(c, x_z, \lambda_z)$ is concave in x_z along the candidate path;

4. the transversality condition holds. That is, for the candidate path, we have

$$\lim_{t\to\infty} e^{-\rho t} x_z(t,b)\lambda_z(t,b) =0.$$

and for all other feasible paths, $\hat{x}_z(t,b)$, we have

$$\lim_{t\to\infty} e^{-\rho t} \hat{x}_z(t,b)\lambda_z(t,b) \geq 0.$$

To prove condition 1, note that starting from $x_z(b,b) = h_z(b)$, we will have $x_z(t,b) \geq 0$ for all $t \geq b$. Moreover, for any path satisfying equations (A3) the flow budget constraint holds:

$$\begin{aligned} r(t)x_z(t,b) - c_z(t,b) =& r(t)x_z(t,b) - \mu(t)x_z(t,b) \\ =& (r(t) - \mu(t))x_z(t,b) \\ =& \dot{x}_z. \end{aligned}$$

To prove condition 2, define $\lambda_z := \mu(t)^{-\sigma} x_z(t,b)^{-\sigma} > 0$. By construction, $c_z(t,b)^{-\sigma} = \lambda_z(t,b)$. Moreover:

$$\begin{aligned} \rho\lambda_z(t,b) - \dot{\lambda}_z(t,b) =& \rho\mu(t)^{-\sigma} x_z(t,b)^{-\sigma} + \mu(t)^{-\sigma}\sigma x_z(s)^{-\sigma-1}\dot{x}_z - \sigma\mu(t)^{-\sigma-1}\dot{\mu}(t)x_z(t,b)^{-\sigma} \\ =& \lambda_z(t,b)\left(\rho + \sigma\frac{\dot{x}_z(t,b)}{x_z(t,b)} + \sigma\frac{\dot{\mu}(t)}{\mu(t)}\right). \end{aligned}$$

Using the equations for $\dot{x}_z(t,b)$ (equation (A3)) and $\dot{\mu}_z(t,b)$ (equation (A22)), we obtain:

$$\rho\lambda_z(t,b) - \dot{\lambda}_z(t,b) = \lambda_z(t,b)\left(\rho + \sigma(r(t) - \mu(t)) + \sigma\left(\mu(t) - r(t) + \frac{1}{\sigma}(r(t) - \rho)\right)\right) = r(t)\lambda_z(t,b)$$

To prove condition 3, note that

$$\max_c H(c, x_z, \lambda_z) = \frac{\lambda_z^{\frac{\sigma-1}{\sigma}}}{1-\sigma} + \lambda_z(r x_z - \lambda_z^{-\frac{1}{\sigma}}),$$

which is concave (linear) in x_z.

To prove the first part of condition 4, note that along the candidate path, x_z grows asymptotically at a rate $\frac{r^*-\rho}{\sigma}$, and λ_z at a rate $\rho - r^*$ ($\mu(t)$ converges along the candidate path). It follows that the first part of the transversality condition holds if

$$-\rho + \frac{r^* - \rho}{\sigma} + \rho - r^* < 0,$$

which is equivalent to the condition $r^* > (r^* - \rho)/\sigma$.

The second part of the transversality condition follows from the fact that, along any feasible path, we have $\hat{x}_z \geq 0$.

It follows that the candidate path provides optimal paths for consumption and asset accumulation outside of the steady state. ■

Proof of Proposition 7

This subsection provides the proof of Proposition 7. Because this proof requires additional material, it is excluded from Appendix A.

Proof of Proposition 7. We start by deriving the equation for $\dot{C}$. Aggregate consumption is given by

$$C(t) = \sum_z \ell_z \int_{-\infty}^t c_z(t,b) p e^{-p(t-b)} \ell_z db + pK(t).$$

Here, $pe^{-p(t-b)}$ is the mass of individuals with skill z from cohort b that are still alive at time t. The term $pK(t)$ is the flow consumption of the dying, which was derived in Lemma A2.

Differentiating this equation, we obtain

$$\begin{aligned}\dot{C}(t) =& \sum_z \ell_z \int_{-\infty}^t \dot{c}_z(t,b) p e^{-p(t-b)} db + p\dot{K}(t) \\ &+ p \sum_z \ell_z c_z(t,t) - p \sum_z \ell_z \int_{-\infty}^t c_z(t,b) p e^{-p(t-b)} db.\end{aligned}$$

These terms can be simplified as follows. The first term captures consumption growth over the life cycle. Using the individual Euler equation, we can rewrite this term as

$$\begin{aligned}\sum_z \ell_z \int_{-\infty}^t \dot{c}_z(t,b) p e^{-p(t-b)} db =& \sum_z \ell_z \int_{-\infty}^t \frac{r(t) - \rho}{\sigma} c_z(t,b) p e^{-p(t-b)} db \\ =& \frac{r(t) - \rho}{\sigma} (C(t) - pK(t)).\end{aligned}$$

The third term denotes the consumption of new cohorts. This is equal to $p\mu(t)X_0(t)$, where $X_0(t)$ denotes the per capita effective wealth of newborns. The fourth term is equal to the reduction in consumption due to individuals entering the dying state. This is equal to $p\mu(t)X(t)$, where $X(t)$ denotes the aggregate effective wealth. Plugging these simplified values in the expression for $\dot{C}$, we obtain

$$\dot{C}(t) = \frac{r(t) - \rho}{\sigma}(C(t) - pK(t)) + p\dot{K}(t) - p\mu(t)(X(t) - X_0(t)).$$

To conclude the derivation of the Euler equation, note that $X(t) - X_0(t) = K(t)$. This follows from the fact that newborns start life with only their human wealth, which is equal to the human wealth in the economy.

The equation for $\dot{K}$ is the usual resource constraint, and the equation for $\dot{\mu}$ was derived above in Lemma A3.

Turning to the distribution of effective wealth, the Kolgomorov Forward Equation (16) follows from the fact that individuals accumulate effective wealth at a rate $r(t) - \mu(t)$ (an implication of Lemma 1) but die with probability p. Upon dying, individuals are replaced by an offspring with no assets, whose effective wealth is therefore $h_z(t)$. ■

C Calibration details

This appendix presents details of the calibration exercise.

Wage data: we compute hourly wages for 1990 using the 1990 Census and for 2014 using the 2012-2016 American Community Survey. We keep the sample of salaried workers between 25 and 54 years of age living in continental US. Following common practice in the literature, we replace top coded wage income by 1.5 times the top code. We compute hourly wages dividing wage income by hours per week times weeks worked. We then converted hourly wages to 2007 dollars using the personal consumption expenditure index, from the BEA. Finally, we winsorized hourly wages between 2 and 180 dollars.

We then computed average wages for workers in each of the 100 wage percentiles. When computing these averages, we weight observations by the Census or ACS weight times total hours of labor supplied, so that we obtain the average hourly wage for workers in each percentile.

Due to changes in the amount of wage data top coded, the data for the top 1 percentile exhibits visible discontinuities over time. We address this issue by fitting a log linear model for log wages as a function of the rank, using the fitted regression to impute the mean wage for the top 1 wage earners. We estimate this model for the percentiles above the 90th percent, exploiting the fact that the top tail of wages has an approximate Pareto shape.

Finally, in Figure 6 (and only for this figure) we smooth the observed change in wages by taking a moving average over consecutive bins 10 percentiles.

Calibration of $\alpha_z(t)$: as explained in the text, we use the year 2000—the midpoint of our period of analysis—to compute ω_{zR}.

Equation (18) can be derived as follows. Let $\bar{\omega}_{zR}$ denote the share of wage income earned by workers with skill z in routine jobs. We assume that $\bar{\omega}_{zR}$ is constant over time but varies across skills. This assumption captures the idea that skill groups have a fixed level of specialization in routine jobs.

Let w_{zR} denote the wage income earned by workers with skill z in routine jobs and let y_{zR} denote the value added generated in those jobs. Let w_{zN} denote the wage income earned by workers with skill z in non-routine jobs and let y_{zN} denote the value added of those jobs.

We have

$$\begin{aligned}\frac{1}{1-\alpha_z(t)} =& \frac{Y_z}{w_z \ell_z}\\ =& \frac{y_{zR} + y_{zN}}{w_{zR} + w_{zN}}\\ =& \frac{w_{zR}}{w_{zR} + w_{zN}} \frac{y_{zR}}{w_{zR}} + \frac{w_{zN}}{w_{zR} + w_{zN}} \frac{y_{zN}}{w_{zN}}\\ =& \bar{\omega}_{zR} \frac{1}{1-\alpha_R(t)} + (1-\bar{\omega}_{zR}) \frac{1}{1-\alpha_0}.\end{aligned}$$

Here $\alpha_R(t)$ denotes the common share of routine tasks that are performed by capital across jobs. α_0 denotes the common share of non-routine tasks that are performed by capital across jobs.

This equation can be rewritten as

$$\frac{1}{1-\alpha_z(t)} = \bar{\omega}_{zR} \left(\frac{1}{1-\alpha_R(t)} - \frac{1}{1-\alpha_0} \right) + \frac{1}{1-\alpha_0}. \tag{A25}$$

Let $\bar{\omega}_z$ denote the share of wage income derived from routine jobs. We also have

$$\begin{aligned}\frac{1}{1-\alpha(t)} =& \frac{\sum_z Y_z}{\sum_z w_z \ell_z}\\ =& \frac{\sum_z y_{zR} + \sum_z y_{zN}}{\sum_z w_{zR} + \sum_z w_{zN}}\\ =& \bar{\omega}_z \frac{1}{1-\alpha_R(t)} + (1-\bar{\omega}_z) \frac{1}{1-\alpha_0}.\end{aligned}$$

This equation can be rewritten as

$$\frac{1}{1-\alpha(t)} - \frac{1}{1-\alpha_0} = \bar{\omega}_z \left(\frac{1}{1-\alpha_R(t)} - \frac{1}{1-\alpha_0} \right). \tag{A26}$$

Combining equations (A25) and (A26), we obtain

$$\frac{1}{1-\alpha_z(t)} = \frac{\bar{\omega}_{zR}}{\bar{\omega}_z}\left(\frac{1}{1-\alpha(t)} - \frac{1}{1-\alpha_0}\right) + \frac{1}{1-\alpha_0}.$$

Equation (18) in the main text follows from the fact that $\omega_{zR} = \frac{\bar{\omega}_{zR}}{\bar{\omega}_z}$.

To operationalize the measurement of $\alpha_z(t)$, we compute ω_{zR} using Census data for 2000. We keep the sample of salaried workers between 25 and 54 years of age living in continental US, and we clean wage data in the exact same way as above. Following the literature, we code an occupation as routine if it is in the top tercile of jobs with the highest routine content according to $O*NET$. We define the routine content of an occupation as total routine inputs minus the average of routine inputs, cognitive inputs, and manual inputs involved in this job. The construction of these inputs is explained in Acemoglu and Autor (2011) and available for download from their websites. We also experimented and obtained similar findings using the classification of occupations as routine and non-routine used in Autor and Dorn (2013). This classification is based on the Dictionary of Occupational Titles, which preceded $O*NET$.

Calibration of parameters of the model: In what follows, we use a superscript d to denote data.

We calibrate the values for γ_z to match 1980 wages. In particular, we have

$$w_z^d(1980) = \gamma_z(1-\alpha_z(1980))Y(1980).$$

It follows that

$$\gamma_z = \frac{w_z^d(1980)/(1-\alpha_z(1980))}{\sum_v w_v^d(1980)/(1-\alpha_v(1980))},$$

which we compute using Census data on average wages by percentile (described above) and our estimates for $\alpha_z(t)$.

We target a capital-output ratio of 3 in 1980. Coupled with the fact that $\alpha(1980) = 0.345$, we obtain

$$R = \alpha\frac{Y}{K} = 0.345 \times (1/3) = 0.115.$$

We compute ψ_z to ensure that automation reduces costs by 30%. This implies

$$\frac{w_z}{\psi_z R} = 1.3 \rightarrow \psi_z = \frac{w_z}{1.3R}.$$

The remaining technology parameter, A is calibrated to match the level of output per hour of labor in 1980.

Turning to the household side, we follow the standard practice of imposing $\sigma = 3$. The value of $R = 11.5\%$ and the assumed depreciation rate of 5% imply $r = 6.5\%$. The remaining parameters, p and ρ, satisfy the equation

$$r = \rho + p\sigma\alpha_{net}(1980),$$

where

$$\alpha_{net}(1980) = \frac{\alpha(1980) - \delta(K/Y)}{1 - \delta(K/Y))} = 23\%.$$

To separate ρ and p, we need an extra equation. We exploit the fact that as p rises (and ρ declines), the capital supply in the model becomes more inelastic. We pick a level of $p = 3.85\%$ which gives a capital-supply elasticity in the 1980 steady state of $d \ln K/dr = 51.28$.

Empirical estimates of the capital-supply elasticity: As mentioned in the main text, an elasticity of capital to rental rates of 50 is large relative to the available empirical estimates. The following table summarizes our review of the empirical literature estimating this elasticity. For each paper, we report the implied elasticity $d \ln K/dr$, which can be directly compared to our calibration target of 50. As is apparent from our survey, all these estimates put the elasticity below 35, which is much more inelastic than what our model predicts.

Study	What Elasticity?	Formula	Estimate	Implied $d \ln K/dr$	Duration	Methodology
Wealth Taxation:						
Zoutman (2015)	Stock of housing and financial assets w.r.t tax on their sum	$\frac{d\log(K)}{d\log(\tau)}$	-0.045	**9**	5 Years	Quasi experiment - 2001 Dutch capital tax reform. Response of households is tracked using panel data.
Jakobsen et al. (2018)	Taxable wealth w.r.t. net-of-tax rate for the very rich / moderately rich	$\frac{d\log(K)}{d\tau}$	[-8,-25]	**[8,25]**	8 years	Quasi experiment - 1989 wealth tax reform in Denmark. Diff-in-diff using variation in tax ceilings and the tax exemption level.
Brülhart et al. (2017)	Semi-elasticity - taxable wealth (excluding pensions) w.r.t. wealth tax rate	$\frac{d\log(K)}{d\tau}$	[-23,-35]	**[23,35]**	-	Cross-regional and time variation in the Swiss wealth-tax system.
Income Taxation:						
Kleven and Schultz (2014)	Positive taxable capital income w.r.t. net-of-tax rate	$\frac{d\log(rK)}{d\log(1-\tau)}$	[0.1,0.3]	**[1.25,7.5]**	3-7 years	Danish tax reforms and full-population administrative data since 1980.

D Generalizations

This appendix presents several generalizations of our baseline model. First, we consider extensions where we explore different ways of breaking the perfect dynasties assumption. We first present a version in which individuals are allowed to purchase annuities, as in Blanchard (1985). We then assume that there is population growth and human capital depreciation, and retain the assumption that individuals consume their wealth upon death. In both cases, we provide steady-state relationships characterizing the return to wealth and the distribution of effective wealth in terms of the net capital share, taxes, and the capital-output ratio (all observables).

Then we move to versions of the model that include profits, taxes on capital income, investment-specific technical change, and sustained economic growth. Variants of these extensions are discussed in Section 3 and at different points in the paper.

Annuities: We deviate from our baseline model in the following ways:

- Assume that there is an annuity market that pays a flow income of $pa_z(s)$, and when individuals die, they give their wealth to the insurance company rather than consume it.

Consumption decisions solve the problem

$$\max_{\{c_z(s),a_z(s)\}_{s\geq 0}} \int_0^\infty e^{-(\varrho+p)s}\frac{c_z(s)^{1-\sigma}}{1-\sigma}ds \tag{A27}$$
$$\text{s.t. } \dot{a}_z(s) = w_z + (r+p)a_z(s) - c_z(s), \text{ and } a_z(s) \geq -w_z/(r+p)$$

Analogous to Lemma 1, the optimal saving and consumption policy functions are

$$\dot{a}_z(s) = \frac{r-\varrho}{\sigma}\left(a_z(s) + \frac{w_z}{r+p}\right), \qquad c_z(s) = \left(r+p-\frac{r-\varrho}{\sigma}\right)\left(a_z(s)+\frac{w_z}{r+p}\right)$$

with $a_z(0)=0$. Let X denote effective wealth of the economy in steady state. We have that $X = H + K$, where $H = \bar{w}/(r+p)$ denotes human wealth, and K denotes the value of the capital stock.

Following the derivation in the main text, it follows that the aggregate behavior of X is given by:

$$0 = \dot{X} = \frac{r-\varrho}{\sigma}X - pK. \tag{A28}$$

Relative to our baseline model, the difference is that in this equation, the rate of accumulation depends on $r-\varrho$, rather than $r-\rho$. This difference is driven by the incentives to accumulate assets introduced by the annuity market: wealth now pays return $r+p$ so that the individual wealth accumulation rate is $(r+p-\varrho-p)/\sigma=(r-\varrho)/\sigma$. As in our baseline model, we still have the term $-pK$ on the right-hand side of equation (A28). This term now captures the wealth paid by dying individuals to the insurance company upon death (not their last-day consumption). The insurance company redistributes this wealth to living households via a higher return to their wealth $r+p$. A market for annuities breaks the perfect dynasties assumption because the wealth of the dying is redistributed to all individuals and not just to newborns (this logic shows that estate taxation will have the same effect as long as revenues are not rebated entirely to newborns).[41]

Note that equation (A28) is equivalent to the equations for aggregates in Blanchard (1985) who analyzes the special case of logarithmic utility, $\sigma=1$. In particular, using our notation, equations (5), (6) and (7) in the published version of his paper are: $C=(\varrho+p)(H+K), \dot{H}=(r+p)H-\bar{w}$ and $\dot{K}=rK+\bar{w}-C$. Using our definition of effective wealth $X:=H+K$, we have $\dot{X}=(r+p)H+rK-(\varrho+p)(H+K)=(r-\varrho)X-pK$ which is the special case of (A28) with $\sigma=1$.

Aggregate capital income now includes annuity payments and so is given by $(r+p)K$. Therefore, rearranging (A28) and using that $X=K+\bar{w}/(r+p)$, the steady state return to wealth satisfies

$$r=\varrho+p\sigma\frac{(r+p)K}{(r+p)K+\bar{w}}=\varrho+p\sigma\alpha_{net}.$$

Alternatively, the analogue to (8) for long-run capital supply is $(K/\bar{w})^s=\frac{r-\varrho}{(\sigma p+\varrho-r)(r+p)}$. Following the same steps as in the baseline model, it follows that effective wealth follows a Pareto distribution. Wages give the scale parameters, and the common tail parameter is given by

$$\frac{1}{\zeta}=\frac{1}{p}\frac{r-\varrho}{\sigma}=\alpha_{net},$$

where we have used that the rate of individual wealth accumulation depends on $r-\varrho$ as discussed below equation (A28).

The model with annuities therefore yields exactly the same expressions for the steady state return to wealth and the tail parameter of the wealth distribution as in our baseline model except for one difference: the mortality-adjusted discount rate $\rho=\varrho+p$ is replaced

[41] Another way of seeing why, also with annuities, there is a term $-pK$ in (A28) is from the law of motion of aggregate capital which is given by $\dot{K}=\bar{w}+(r+p)K-pK-C=\bar{w}+rK-C$. As Blanchard (1985) explains: "Whereas individual wealth accumulates, for those alive, at rate $r+p$, aggregate wealth accumulates at rate r. This is because the amount pK is a transfer, through life insurance companies, from those who die to those who remain alive; it is not therefore an addition to aggregate wealth." Integrating across individual consumption policy functions, aggregate consumption is $C=\left(r+p-\frac{r-\varrho}{\sigma}\right)(K+H)$. Substituting in, we have $\dot{K}=\bar{w}+(r+p)K-pK-C=\frac{r-\varrho}{\sigma}(K+H)-pK$ which is equivalent to (A28).

by the unadjusted discount rate ϱ.

Population growth and human capital depreciation: We deviate from our baseline model in the following ways:

- Assume that population grows at a rate n. That is, the p individuals who dies ta time t are replaced by $p+n$ individuals.
- Assume that as individuals age, their human capital depreciates at a rate δ_H. To simplify the algebra. Assume that the mass of individuals with skill z is $\ell_z \frac{p+n+\delta_H}{p+n}$, so that the total supply of skill z labor is ℓ_z.

Consumption decisions solve the problem

$$\max_{\{c_z(s), a_z(s)\}_{s \geq 0}} \int_0^\infty e^{-(\varrho+p)s} \frac{c_z(s)^{1-\sigma}}{1-\sigma} ds \tag{A29}$$
$$\text{s.t. } \dot{a}_z(s) = w_z e^{-\delta_H s} + r a_z(s) - c_z(s), \text{ and } a_z(s) \geq -w_z/(r+\delta_H)$$

Let X denote effective wealth of the economy in steady state. We have that $X = H + K$, where $H = \bar{w}/(r+\delta_H)$ denotes human wealth, and K denotes the value of the capital stock.

Following the derivation in the main text, it follows that the aggregate behavior of X is given by:

$$nX = \dot{X} = \frac{r-\rho}{\sigma} X - pK + (n+\delta_H)H.$$

Relative to our baseline case, there are two differences in this equation. First, total wealth grows at a rate n since population increases over time. Second, the term $(n+\delta_H)H$ captures the fact that newborns have more human wealth due to their sheer number but also the fact that human capital depreciates as individuals age.

Rearranging this equation, we obtain:

$$\begin{aligned} r &= \rho - \delta_H \sigma + (p+n+\delta_H)\sigma \frac{(r+\delta_H)K}{(r+\delta_H)K + \bar{w}} \\ &= \rho - \delta_H \sigma + (p+n+\delta_H)\sigma \frac{\alpha_{net} + \delta_H K/Y_{net}}{1 + \delta_H K/Y_{net}}. \end{aligned}$$

Let $g_z(x)$ denote the distribution of x_z. Following the same steps as in the baseline model, we obtain the Kolgomorov Forward Equation for $\tilde{x}_z$:

$$0 = -\left(\frac{r-\rho-\sigma g}{\sigma} x g_z(x)\right)' - p g_z(x) - n g_z(x)$$

on $(w_z/r, \infty)$. The extra term $-ng_z(x)$ captures the loss of probability mass as new individuals are born and start their lives with effective wealth w_z/r.

Following the same steps as in the baseline model, it follows that effective wealth follows a Pareto distribution. Wages give the scale parameters, and the common tail parameter is given by

$$\frac{1}{\zeta} = \frac{p+n+\delta_H}{p+n} \frac{\alpha_{net} + \delta_H K/Y_{net}}{1+\delta_H K/Y_{net}} - \frac{\delta_H}{p+n}$$

Profits: We deviate from our baseline model in the following ways:

- Assume that producers of the final good obtain profits and that profits are equal to a share $\pi \in [0,1)$ of net output.

Let X denote effective wealth of the economy in steady state. We have that $X = H + K + V$, where $H = \bar{w}/r$ denotes human wealth, $V = \pi Y_{net}/r$ denotes the equity value of the firm, and K denotes the value of the capital stock. Here $Y_{net} = Y - \delta K$ denotes net output.

Following the derivation in the main text, it follows that the aggregate behavior of X is given by:

$$0 = \dot{X} = \frac{r-\rho}{\sigma} X - p(K+V).$$

The last term captures the difference in effective wealth between newborns and existing cohorts, which coincides with the financial wealth in the economy.

Rearranging this equation, we obtain:

$$r = \rho + p\sigma \frac{rK + \pi Y_{net}}{rK + \pi Y_{net} + \bar{w}} = \rho + p\sigma(\alpha_{net}(1-\pi) + \pi)$$

The term $\alpha_{net}(1-\pi)+\pi$ has a simple interpretation: it gives the share of net capital income (inclusive of profits) in net income. This equals one minus the share of labor in net income.

Following the same steps as in the baseline model, it follows that effective wealth follows a Pareto distribution. Wages give the scale parameters, and the common tail parameter is given by

$$\frac{1}{\zeta} = \alpha_{net}(1-\pi) + \pi.$$

Taxes and investment-specific technical change: We deviate from our baseline model in the following ways:

- Assume that capital income (including profits) is taxed at a rate $\tau_K \in [0,1)$. The revenue is used to finance a public good that does not distort saving and consumption decisions.

- Assume that capital can be produced using q_K units of the final good.

Let X denote effective wealth of the economy in steady state. We have that $X = H + (1/q_K)K$, where $H = \bar{w}/r$ denotes human wealth, and $(1/q_K)K$ denotes the value of the capital stock (here we use K to denote the quantity of capital).

Following the derivation in the main text, it follows that the aggregate behavior of X is given by:

$$0 = \dot{X} = \frac{r-\rho}{\sigma}X - p(1/q_K)K.$$

The last term captures the difference in effective wealth between newborns and existing cohorts, which coincides with the financial wealth in the economy.

Rearranging this equation, we obtain:

$$r = \rho + p\sigma\frac{r(1/q_K)K}{r(1/q_K)K + \bar{w}} = \rho + p\sigma\frac{(1-\tau_K)\alpha_{net}}{(1-\tau_K)\alpha_{net} + 1 - \alpha_{net}}.$$

The term $(1-\tau_K)\alpha_{net}/((1-\tau_K)\alpha_{net} + \alpha_{net})$ has a simple interpretation: it gives the share of net after-tax capital income in net after-tax income. This is equal to one minus the share of labor in after tax income.

Following the same steps as in the baseline model, it follows that effective wealth follows a Pareto distribution. Wages give the scale parameters, and the common tail parameter is given by

$$\frac{1}{\zeta} = \frac{(1-\tau_K)\alpha_{net}}{(1-\tau_K)\alpha_{net} + \alpha_{net}}.$$

Economic growth: We deviate from our baseline model in the following ways:

- Assume that ψ_z grows at a constant rate g, so that in steady state the economy also grows at this rate.

Let X denote effective wealth of the economy in steady state. We have that $X = H + K$, where $H = \bar{w}/r$ denotes human wealth and K denotes the value of the capital stock.

Following the derivation in the main text, it follows that the aggregate behavior of X is given by:

$$gX = \dot{X} = \frac{r-\rho}{\sigma}X - pK.$$

The only difference is that now effective wealth grows at a rate g in steady state.

Rearranging this equation, we obtain:

$$r = \rho + \sigma g + p\sigma \frac{rK}{rK + \bar{w}} = \rho + \sigma g + p\sigma \alpha_{net}$$

We now turn to the distribution of $\tilde{x}_z = x_z/Y$. The normalization ensures that the distribution converges. Otherwise, the reinjection points w_z/r grow at a rate g.

Let $g_z(x)$ denote the distribution of $\tilde{x}_z$. As individuals age, $\tilde{x}_z$ grows at a constant rate

$$\frac{\dot{\tilde{x}}_z}{\tilde{x}_z} = \frac{\dot{x}_z}{x_z} - g = \frac{r - \rho - \sigma g}{\sigma}.$$

Following the same steps as in the baseline model, we obtain the Kolgomorov Forward Equation for $\tilde{x}_z$:

$$0 = -\left(\frac{r - \rho - \sigma g}{\sigma} x g_z(x)\right)' - p g_z(x)$$

on $(w_z/(rY), \infty)$. Note that in steady state, the reinjection points $w_z/(rY)$ are constant, ensuring that the distribution of $\tilde{x}_z$ converges.

This implies that x_z follows a Pareto distribution. Wages give the scale parameters (and hence shift over time as the economy grows). The common tail parameter is given by

$$\frac{1}{\zeta} = \frac{r - \rho - \sigma g}{p\sigma} = \alpha_{net}.$$

E Appendix for Section 3

A.1 Returns to Private Equity

[TO BE ADDED]

A.2 Decomposing Percentile-Specific Income Growth

Average income $y_t(q)$ in a given percentile q and year t is the sum of average labor income $y_{\ell,t}(q)$ and average capital income $y_{k,t}(q)$:

$$y_t(q) = y_{\ell,t}(q) + y_{k,t}(q).$$

We want to decompose the growth rate over a T-year time period relative to a base year $t = 0$

$$g_T(q) = \frac{y_T(q) - y_0(q)}{y_0(q)}$$

Our goal is to derive a decomposition of $y(q)$ and $g(q)$ that answers the question: what would the p-th percentile income have been if capital income had stayed at its base level (say in 1980)? The answer is

$$y_{\ell,T}(q) + y_{k,0}(q)$$

We therefore decompose the T-year growth rate as

$$g_T(q) = \frac{y_T(q) - y_0(q)}{y_0(q)} = \underbrace{\frac{y_{\ell,T}(q) - y_{\ell,0}(q)}{y_0(q)}}_{\text{part due to labor inc}} + \underbrace{\frac{y_{k,T}(q) - y_{k,0}(q)}{y_0(q)}}_{\text{part due to capital inc}}. \tag{A30}$$

In Figure 13, we plot the *annualized* T-year growth rate relative to a base year $t = 0$

$$g(q) = \left(\frac{y_T(q)}{y_0(q)}\right)^{1/T} - 1$$

and want to decompose it in an analogous fashion. To achieve this, we simply keep track of the *shares* in the decomposition (A30) and then scale the annualized T-year growth rate $g(q)$ by those shares to compute an annualized decomposition. That is, denote the share of growth due to labor income by $\omega_\ell(q) = \frac{(y_{\ell,T}(q) - y_{\ell,0}(q))/y_0(q)}{(y_T(q) - y_0(q))/y_0(q)}$. Then the part of annualized growth due to labor is $g_\ell(q) = \omega_\ell(q)g(q)$ and the part due to capital income is $g_k(q) = (1 - \omega_\ell(q))g(q)$.

An alternative strategy for decomposing the annualized growth rate $g(q)$ is to define the part of income growth due to labor income as $\widehat{g}_\ell(q) = \left(\frac{y_{\ell,T}(q) + y_{k,0}(q)}{y_0(q)}\right)^{1/T} - 1$ and the part due to capital income as $\widehat{g}_k(q) = \left(\frac{y_{\ell,0}(q) + y_{k,T}(q)}{y_0(q)}\right)^{1/T} - 1$. The disadvantage of this alternative decomposition is that, due to compound growth, it is not additive. We therefore prefer scaling the annualized growth rate by the shares $\omega_\ell(q)$ and $\omega_k(q)$, as discussed in the preceding paragraphs.