Online Additional Material

1 Extensions

This section discusses the extensions described in Section 4.4. We first introduce "learning-from-others" by allowing the innate ability distribution $\bar{s}_t(i)$ to depend on the economy’s skill distribution $s_t(i)$. Second, we introduce re-training by allowing a fraction of the older generations to make skill investment following the shock. Finally, we introduce population growth by allowing the birth rate to be higher than the death rate.

1.1 Learning-from-others

We relax the assumption that the reference distribution $\bar{s}_t(i)$ in the skill investment problem is exogenous and fixed over time. Instead, we assume that certain skills may be easier to acquire than others because workers can "learn from others" when such skills are already abundant in the economy. Formally, we assume that the baseline distribution $\bar{s}_t(i)$ is a geometric average of a fixed distribution $\bar{\epsilon}(i)$ and the current skill distribution in the economy $s_t(i)$ at the time where generation $\tau$ is born,

$$\bar{s}_t(i) = s_t(i)^\gamma \bar{\epsilon}(i)^{1-\gamma}, \quad \gamma \in [0, 1).$$

Note that as $\gamma$ increases it becomes easier for workers to choose skill lotteries that put more weight on those skill types that are already abundant in the economy. As opposed to our benchmark case ($\gamma = 0$), this extension with $\gamma > 0$ introduces a backward-looking element to the skill investment problem and complementarities in skill investment decisions across generations.

In what follows, we reproduce the key steps that change in the proofs in Appendix A.3. First, we log-linearize the extended version of (A.6). We begin by noting that the stationary distribution exist and is

$$s(i) = \frac{s(i)^\gamma \epsilon(i)^{1-\gamma} w(i)^{\psi}}{\int_0^1 s(j)^\gamma \epsilon(j)^{1-\gamma} w(j)^{\psi} dj} \implies s(i) = \frac{\epsilon(i) w(i)^{1-\gamma} \psi}{\int_0^1 \epsilon(j) w(j)^{1-\gamma} \psi dj}.$$

Then, we obtain that

$$\hat{s}_t(i) = \gamma (\bar{s}_t(i) - \bar{s}_t(l)) + \hat{s}_t(l) - \psi \hat{q}_t \mathbb{I}_{i < t} - \psi \hat{q}_{t+\tau} \mathbb{I}_{i \in (t, t)}.$$

Second, we replace the above in the expression inside the parenthesis in (A.4), we obtain
\[
\gamma \int_{l_t}^{1} \frac{\alpha(i)\sigma(i)s(i)}{\int_{l_t}^{1} \alpha(x)\sigma(x)s(x)dx} di - \int_{0}^{l_t} \frac{\alpha(i)s(i)}{\int_{0}^{l_t} \alpha(x)s(x)dx} di = \gamma \beta \int_{l_t}^{1} \frac{\alpha(i)\sigma(i)s(i)}{\int_{l_t}^{1} \alpha(x)\sigma(x)s(x)dx} di - \int_{0}^{l_t} \frac{\alpha(i)s(i)}{\int_{0}^{l_t} \alpha(x)s(x)dx} di
\]

where the last line uses (A.3) and (A.2).

Third, as in the proof in Appendix A.3, we can show that the last term inside the integral is of second order. Thus, replacing the above expression back in (A.4), we obtain the Kolmogorov-Forward equation for \( \hat{l}_t \) in the economy with learning-from-others,

\[
\frac{\partial \hat{l}_t}{\partial t} = -\delta (1 - \gamma) \hat{l}_t + \frac{\eta}{\kappa \theta + \delta} \delta \hat{q}_t.
\]

Fourth, since the law of motion for \( \hat{q}_t \) is the same as in the benchmark model, this implies that the equilibrium is saddle-path stable where the new \( \lambda \) in the economy with learning-from-others is the positive solution to

\[
(\delta (1 - \gamma) - \lambda)(\rho + \delta + \lambda) + \frac{\psi \delta}{\kappa \theta + \delta} = 0.
\]

Finally, the optimal lottery in the economy with learning-from-others is

\[
\hat{s}_t(i) = \gamma \hat{s}_t(i) + \left( \mathbb{1}_{i > l_t} - \int_{l_t}^{1} s(i)di \right) \psi \hat{q}_t + \omega_t(i).
\]

Next, we reproduce the key steps that change in Appendices A.4 and A.7. First, from the expression for the stationary distribution above, note that the long-run skill supply elasticity in the learning-from-others economy is \( \frac{1}{1 - \gamma} \psi \) as opposed to simply \( \psi \).

This implies that the dynamic responses are

\[
\frac{\partial \hat{l}_t}{\partial t} = -\delta (1 - \gamma) \hat{l}_t + \frac{\eta}{\kappa \theta + \delta} \delta \hat{q}_t.
\]
\[ \Delta \log(l_t) = -\frac{\eta}{\theta + \kappa \eta} \left( 1 + \frac{1}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta}} \frac{1}{1 - \gamma \rho + \delta} (e^{-\lambda t} - 1) \right) (\theta - 1) \Delta \log(A) \]

\[ \Delta \log(y_t) = \frac{1}{\theta + \kappa \eta} \left( (1 + \kappa \eta) + \frac{(\theta - 1)}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta}} \frac{1}{1 - \gamma \rho + \delta} (1 - e^{-\lambda t}) \right) (\theta - 1) \Delta \log(A) \]

\[ \Delta \log(q_t) = \frac{1}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta}} \left( 1 + \frac{\lambda - \delta(1 - \gamma)}{\delta(1 - \gamma)} e^{-\lambda t} \right) (\theta - 1) \Delta \log(A) \]

where the last line follows from the equation for the new \( \lambda \).

Second, note that the short-run responses for \( l_t \) and \( y_t \) are identical than in the benchmark model. The long-run responses are larger (smaller) in magnitude for \( y_t \) (for \( l_t \)) in the economy with learning-from-others since the long-run skill supply elasticity is larger and thus \( \frac{1}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta}} \frac{1}{1 - \gamma \rho + \delta} \) is larger. As for the DCIR, note that \( \lambda \) is smaller in the learning-from-others economy. Together with the fact that \( \frac{1}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta}} \frac{1}{1 - \gamma \rho + \delta} \) is larger, they imply that the DCIR of both \( y_t \) and \( l_t \) is higher in the learning-from-others economy.

Third, for \( q_t \) we have that

\[ \Delta \log(q_\infty) = \frac{1}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta}} \left( 1 \right) (\theta - 1) \Delta \log(A) \]

\[ \Delta \log(q_0) = \frac{1}{\theta + \kappa \eta} \left( \rho + \delta + \frac{1}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta + \lambda}} \right) (\theta - 1) \Delta \log(A) \]

\[ \int_0^\infty \hat{q}_t dt = \frac{1}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta}} \frac{1}{\theta + \kappa \eta + \frac{1}{1 - \gamma \rho + \delta + \lambda}} \left( \frac{1}{1 - \gamma} \right) (\theta - 1) \Delta \log(A). \]

Then, since \( \lambda \) is smaller, the short- and long-run responses are smaller in magnitude and the DCIR is larger in the economy with learning-from-others.

Finally, we note that the proofs for the comparative statics in Appendix A.7 with respect to \( \eta \) and \( \psi \) are unchanged. To see this, it suffices to show that the dynamics for \( q_t, l_t, y_t \) in the economy with learning-from-others are equivalent to those from a re-parameterized benchmark economy where \( \delta' = \delta(1 - \gamma), \psi' = \frac{1}{1 - \gamma} \psi \) and \( \rho' = \rho + \delta \gamma \).
1.2 Old generations skill investment

We now let a fraction of workers that were present before the shock re-optimize their skill investment "as if" they were a young generation entering at time $t = 0$. Formally, the skill distribution on impact now becomes

$$s_0(i) = (1 - \beta)s_{0-}(i) + \beta \tilde{s}_0(i),$$

where $\beta$ is the fraction of workers in the generation present before the shock that can re-optimize.

The first thing to note is that this does not change any of the transitional dynamics given the new initial skill distribution on impact. As such Theorem 1 is unchanged. However, the initial conditions and the dynamic responses do change. Next, we reproduce the key steps that change in Appendix A.4.

The deviation from the skill distribution on impact from the new stationary distribution is now

$$\hat{s}_0(i) = \hat{s}_{0-}(i) + \beta (\hat{s}_0(i) - \hat{s}_{0-}(i))$$

where the long-run change $\Delta \log(\omega)$ is the same as in the benchmark model.

Following the same steps as in the benchmark proof, this then implies that

$$\left(\frac{\theta}{\eta} + \kappa\right) \hat{l}_0 = \int_1^l \frac{\sigma(i)\alpha(i)s(i)}{\int_1^l \sigma(i)\alpha(i)s(i)} \hat{s}_0(i) di - \int_1^l \frac{\alpha(i)s(i)}{\int_0^l \alpha(i)s(i)} \hat{s}_0(i) di$$

$$= - (1 - \beta) \frac{\psi}{\rho + \delta} \Delta \log(\omega) + \beta \psi \hat{q}_0.$$

Thus,

$$\hat{\omega}_0 = - \frac{1}{\eta} \hat{l}_0$$

$$= \frac{1}{\theta + \kappa \eta} \left( \frac{\psi}{\rho + \delta} \Delta \log(\omega) - \beta \left( \frac{\psi}{\rho + \delta} \Delta \log(\omega) + \psi \hat{q}_0 \right) \right)$$

$$= \frac{1}{\theta + \kappa \eta} \left( \frac{\psi}{\rho + \delta} \Delta \log(\omega) - \beta \left( \frac{\psi}{\rho + \delta} \Delta \log(\omega) + \frac{\psi}{\rho + \delta + \lambda} \hat{\omega}_0 \right) \right)$$

$$= \frac{1 - \beta}{1 + \beta \frac{\psi}{\rho + \delta + \lambda} \frac{1}{\theta + \kappa \eta} \frac{1}{\rho + \delta}} \frac{\psi}{\theta + \kappa \eta} \Delta \log(\omega).$$

Finally, using the above together with the expression for $\Delta \log(\omega)$ in equations
(A.8)-(A.10), we obtain:

\[
\Delta \log(y_t) = \frac{1}{\theta + \kappa \eta} \left( 1 + \kappa \eta + (\theta - 1) \frac{\psi}{\lambda} \left( 1 - \frac{1 - \beta}{1 + \beta \frac{\lambda - \delta}{\delta}} e^{-\lambda t} \right) \right) (\theta - 1) \Delta \log(A)
\]

\[
\Delta \log(q_t) = \frac{1}{\chi} \left( 1 + \frac{\lambda - \delta}{\delta} \frac{1 - \beta}{1 + \beta \frac{\lambda - \delta}{\delta}} e^{-\lambda t} \right) (\theta - 1) \Delta \log(A)
\]

\[
\Delta \log(l_t) = -\frac{\eta}{\theta + \kappa \eta} \left( 1 + \frac{\psi}{\lambda} \left( \frac{1 - \beta}{1 + \beta \frac{\lambda - \delta}{\delta}} e^{-\lambda t} - 1 \right) \right) (\theta - 1) \Delta \log(A)
\]

Then, mathematically, the dynamic responses in the economy where old generations can re-optimize their skills are similar to those in the benchmark economy except that the function \(e^{-\lambda t}\) is now multiplied by \(\frac{1 - \beta}{1 + \beta \frac{\lambda - \delta}{\delta}} < 1\). This immediately implies that: the long-run responses are the same in both economies, the short-run responses of \(y\) and \(l\) (of \(q\)) are now larger (smaller) in magnitude, and the DCIR of all variables is now smaller. Hence, in many ways, this new economy behaves qualitatively similar to an economy with a lower degree of skill specificity (higher \(\eta\)), with the exception that long-run responses are unchanged.

### 1.3 Population growth

We now assume that the size of entering generations is \(\mu\) as opposed to \(\delta\). This implies that the population growth rate is \(\mu - \delta\). The Kolmogorov-Forward equation describing the evolution of the skill distribution becomes

\[
\frac{\partial e^{(\mu - \delta)t} s_t(i)}{\partial t} = -\delta e^{(\mu - \delta)t} s_t(i) + \mu e^{(\mu - \delta)t} \tilde{s}_t(i).
\]

Then, we have that

\[
\frac{\partial s_t(i)}{\partial t} = -\mu s_t(i) + \mu \tilde{s}_t(i).
\]

The remaining elements in the model remain the same. Hence, the economy with population growth is identical to our benchmark economy except that the convergence rate \(\lambda\) is higher iff \(\mu > \delta\) since it is now the positive solution to:

\[
(\lambda - \mu)(\rho + \delta + \lambda) = \frac{\psi \mu}{\theta + \kappa \eta}.
\]

Then, if \(\mu > \delta\), the short- and long-run dynamic responses for \(y_t, l_t\) remain unchanged, the short-run response of \(q\) is smaller in magnitude, and the DCIR of
all variables is lower. The opposite holds when $\mu < \delta$. 
2 Numerical Analysis

This appendix discusses in detail the parameterization of the model. We first present the theoretical impulse response functions for the relative employment of different worker generations. Second, we describe the procedure to select the parameters that match the theoretical and empirical impulse response functions. Finally, we use the parametrized model to quantitatively evaluate the dynamic adjustment to cognitive-biased technological innovations.

2.1 Impulse response functions of relative employment by generation

As a first step to parametrize our theory using the empirical impulse response functions in Section 6, we derive the theoretical responses of generation-specific relative employment. To this end, we consider the same one-time permanent change in $A$ at $t = 0$. We define older generations as those born before period $t = -x$ and younger generations as those born at period $t = -x$. In period $t \geq 0$, the relative high-tech employment of these worker generations are given by

$$e_{t}^{\text{old}} = \frac{\int_{l_{t}}^{1} s_{0}(i)di}{\int_{l_{t}}^{1} s_{0}(i)di}$$

and

$$e_{t}^{\text{young}} = \frac{\tilde{x}_{0}e^{-\delta t} \int_{l_{t}}^{1} s_{0}(i)di + \delta \int_{0}^{t} e^{\delta(\tau-t)} \int_{l_{t}}^{1} \tilde{s}_{\tau}(i)di d\tau}{\tilde{x}_{0}e^{-\delta t} \int_{0}^{t} s_{0}(i)di + \delta \int_{0}^{t} e^{\delta(\tau-t)} \int_{0}^{t} \tilde{s}_{\tau}(i)di d\tau},$$

where $\tilde{x}_{0} \equiv 1 - e^{-\delta x}$ is the population share of the young generation at $t = 0$.

For both worker groups, the technology-skill assignment is identical and determined by the threshold $l_{t}$. Notice that all workers of the old generations have the pre-shock skill distribution, $s_{0}(i)$. However, the skill distribution of young generations combines the pre-shock distribution, $s_{0}(i)$, and the post-shock lotteries, $\tilde{s}_{\tau}(i)$. The overlapping generation structure of the model implies that the relative share of workers in the young generation with the pre-shock skill distribution decays at the constant rate $\delta$.

We allow the young group to include workers born before the shock (since $x \geq 0$). This circumvents the challenge of identifying the cohorts that start adjusting their skills after the shock, which arises because, in practice, technologies may not be adopted instantaneously and young workers may still invest on skills after entering the labor force (in the form of vocational training or on-the-job learning). It is also possible to allow part of the workers born before the shock to adjust their skills at $t = 0$. In this case, rather than $s_{0}(i)$, the initial skill distribution would be a mix of $s_{0}(i)$ and $\tilde{s}_{0}(i)$. This extension does not alter our main qualitative insights, but reduces the magnitude of the short-to-long adjustment in the economy.
**Relative employment of old generation.** We show below that the change in the relative employment of old generations is

\[
\Delta \log e_{t}^{\text{old}} \approx \frac{\eta}{\theta + \kappa \eta} e_{H}^{-1} \left(1 - \frac{\psi}{\chi} (1 - e^{-\lambda t}) \right) (\theta - 1) \Delta \log A,
\]

(1)

where \(e_{H}\) is the high-tech employment share at \(t = 0^{-}\).

Among old generations, the increase in the relative productivity of high-tech production induces the reallocation of workers towards high-tech production whenever \(\theta > 1\). The expression indicates that this positive effect on relative high-tech employment becomes weaker over time. This follows from the expansion of high-i skills among younger generations, which displaces old workers with marginal skills from high-tech production – i.e., those with skills \(i \in (l_{0}, l_{\infty})\). Importantly, expression (1) shows that the magnitude of the increase in relative employment of older generations is decreasing in the degree of technology-skill specificity (i.e., increasing in \(\eta\)).

**Relative employment of young generation.** Turning to the employment response among young generations, we show below that

\[
\Delta \log e_{t}^{\text{young}} \approx \Delta \log e_{t}^{\text{old}} + \psi \frac{1 - e^{-\lambda t}}{\chi \left(1 - (1 - \tilde{x}_{0}) e^{-\delta t}\right)} (\theta - 1) \Delta \log A.
\]

(2)

This expression indicates that the evolution of the allocation of young workers has two components. The first term captures the change in technology-skill assignment and, since it is the only determinant of the relative employment of old generations, it can be approximated by \(\Delta \log e_{t}^{\text{old}}\). The second term captures the change in the skill investment decision of incoming cohorts. At each point in time, this term is positive as young workers distort skill investment towards high-i skills that became more valuable in high-tech production. We can also show that the between-generation difference grows shortly after the shock. Importantly, expression (1) indicates that the between-generation difference in the response of relative employment is decreasing in the skill investment cost (i.e., it is increasing in \(\psi\)).

### 2.1.1 Proof of equations (1)–(2)

**Proof of equation (1).** We first use a first-order approximation to write the log-change in relative high-tech employment in terms of changes in the high-tech employment share:

\[
\Delta \log \left( e_{t}^{\text{old}} \right) = \log \left( \frac{e_{t}^{\text{old}}}{e_{0^{-}}^{\text{old}}} \right) \approx \frac{1}{(1 - e_{H,\infty}) e_{H,\infty}} \left( e_{H, t}^{\text{old}} - e_{H, 0^{-}}^{\text{old}} \right)
\]

100
where \( e_{H,t}^{\text{old}} = \int_l^1 s_0(i) \, di \).

Since \( \Delta \left( \frac{1}{(1-e_{H,\infty})e_{H,0^-}} \right) \left( e_{H,t}^{\text{old}} - e_{H,0^-}^{\text{old}} \right) \) is a second order term, we get the approximation:

\[
\Delta \log \left( e_t^{\text{old}} \right) \approx \frac{1}{(1-e_{H,0^-})e_{H,0^-}} \left( e_{H,t}^{\text{old}} - e_{H,0^-}^{\text{old}} \right)
\]

We have that

\[
e_{H,t}^{\text{old}} - e_{H,0^-}^{\text{old}} = \int_l^1 s_0(i) \, di - \int_{l_0^-}^1 s_0(i) \, di
\]

By approximating these expressions around \( l \),

\[
e_{H,t}^{\text{old}} - e_{H,0^-}^{\text{old}} \approx -s_0(1) l \left( \Delta \log(l_{\infty}) + \hat{\ell}_t \right)
\approx (s_0(1) l) \eta \Delta \log(\omega_t)
\approx (s_0(l_0^-) l_0^-) \eta \Delta \log(\omega_t)
\approx (1 - e_{H,0^-}) \eta \Delta \log(\omega_t)
\]

where the third equality follows from the fact that \( \Delta (s_0(l) l) \Delta \log(\omega_t) \) is a second order term, and the last equality follows from normalizing the initial skill distribution to be uniform (which implies \( s_0(l_0^-) l_0^- = 1 - e_{H,0^-} \)).

Combining the two expressions,

\[
\Delta \log \left( e_t^{\text{old}} \right) \approx \frac{1}{e_{H,0^-}} \eta \Delta \log(\omega_t)
\]

Using the demand expression in (2),

\[
\Delta \log \left( e_t^{\text{old}} \right) \approx \frac{1}{e_{H,0^-}} \eta \left( -\frac{1}{\theta - 1} \log y_t + \Delta \log A \right)
\]

Using the expression for the evolution of \( y_t \) in Proposition 1,

\[
\Delta \log \left( e_t^{\text{old}} \right) \approx \frac{1}{e_{H,0^-}} \frac{\eta}{\theta + \kappa \eta} \left( -1 - \kappa \eta - \frac{\psi}{\chi} (\theta - 1) (1 - e^{-\lambda t}) + (\theta + \kappa \eta) \right) \Delta \log A
\]

which is identical to (1).

**Proof of equation (2).** We first use a first-order approximation to write the log-change in relative high-tech employment in terms of changes in the high-tech
employment share:

$$\log \left( \frac{e_{H,t}^{\text{young}}}{e_{0-}^{\text{young}}} \right) - \log \left( \frac{e_{H,t}^{\text{old}}}{e_{0-}^{\text{old}}} \right) \approx \frac{1}{1 - e_{H,\infty}} \left( \frac{e_{H,t}^{\text{young}} - e_{H,0}^{\text{young}}}{e_{H,\infty}} - \frac{e_{H,t}^{\text{old}} - e_{H,0}^{\text{old}}}{e_{H,\infty}} \right)$$

$$= \frac{1}{(1 - e_{H,\infty})e_{H,\infty}} \left( \left( \frac{e_{H,t}^{\text{young}}}{e_{H,\infty}} - \frac{e_{H,t}^{\text{old}}}{e_{H,\infty}} \right) - \left( \frac{e_{H,0}^{\text{young}}}{e_{H,\infty}} - \frac{e_{H,0}^{\text{old}}}{e_{H,\infty}} \right) \right)$$

where the last equality follows from the fact that before the shock old and young make identical choices, $e_{H,0}^{\text{young}} = e_{H,0}^{\text{old}}$.

Using the definition of employment shares for each generation,

$$e_{H,t}^{\text{young}} - e_{H,t}^{\text{old}} \approx \frac{1}{1 - (1 - \bar{x})e^{-\delta t}} \left( \tilde{x}e^{-\delta t} \int_{l_t}^{1} s_0(i) di + \delta \int_0^t e^{\delta (\tau - t)} \int_{l_t}^{1} \tilde{s}_\tau(i) d\tau d\tau \right) - \int_{l_t}^{1} s_0(i) di$$

$$\approx \frac{1}{1 - (1 - \bar{x})e^{-\delta t}} \left( \delta \int_0^t e^{\delta (\tau - t)} \int_{l_t}^{1} \tilde{s}_\tau(i) - s_0(i) \right) d\tau$$

Thus,

$$\log \left( \frac{e_{H,t}^{\text{young}}}{e_{0-}^{\text{young}}} \right) - \log \left( \frac{e_{H,t}^{\text{old}}}{e_{0-}^{\text{old}}} \right) \approx \frac{1}{(1 - e_{H,\infty})e_{H,\infty}} \frac{1}{1 - (1 - \bar{x})e^{-\delta t}} \left( \delta \int_0^t e^{\delta (\tau - t)} \int_{l_t}^{1} \tilde{s}_\tau(i) - s_0(i) \right) d\tau$$

We now consider the following approximation:

$$\int_{l_t}^{1} (\tilde{s}_\tau(i) - s_0(i)) di \approx \int_{l_t}^{1} s(i) (\tilde{s}_\tau(i) - s_0(i)) di$$

Then, we derive $s_0(i)$ using the expression for the stationary skill distribution

$$s_0(i) = \frac{\tilde{s}(i) \alpha(i) \frac{\psi}{\rho + \delta} (\omega_0 - \sigma(i)) \frac{\psi}{\rho + \delta} \mathbb{I}_{l_0} - \int_{l_0}^{1 - \tilde{s}(j) \alpha(j) \frac{\psi}{\rho + \delta} dj + \int_{l_0}^{1} \tilde{s}(j) \alpha(j) \frac{\psi}{\rho + \delta} (\omega_0 - \sigma(j)) \frac{\psi}{\rho + \delta} dj}}{\int_{l_0}^{1 - \tilde{s}(j) \alpha(j) \frac{\psi}{\rho + \delta} dj + \int_{l_0}^{1} \tilde{s}(j) \alpha(j) \frac{\psi}{\rho + \delta} (\omega_0 - \sigma(j)) \frac{\psi}{\rho + \delta} dj}}$$

$$\implies \tilde{s}_0(i) \approx - \left( \mathbb{I}_{l_0} - \int_{l}^{1} s(j) \right) \frac{\psi}{\rho + \delta} \Delta \log (\omega)$$

Using the third part of Theorem 1,

$$\int_{l_t}^{1} (\tilde{s}_\tau(i) - s_0(i)) di \approx e_{H,\infty} (1 - e_{H,\infty}) \left( \psi \tilde{q}_\tau + \frac{\psi}{\rho + \delta} \Delta \log (\omega) \right)$$

$$= e_{H,\infty} (1 - e_{H,\infty}) \psi (\tilde{q}_\tau + \Delta \log (q))$$

102
We now apply this expression into (3):

\[
\log \left( \frac{e^{\text{young}}_{0}}{e^{\text{old}}_{0}} \right) \approx \frac{\psi}{1 - (1 - \tilde{x}_0)e^{-\delta t}} \left( \delta \int_0^t e^{\delta (\tau - t)} (\dot{q}_\tau + \Delta \log(q)) \, d\tau \right)
\]

\[
\approx \frac{\psi}{1 - (1 - \tilde{x}_0)e^{-\delta t}} \left( \delta \int_0^t e^{\delta (\tau - t)} \dot{q}_0 e^{-\lambda \tau} d\tau + (1 - e^{-\delta t})\Delta \log(q) \right)
\]

\[
\approx \frac{\psi}{1 - (1 - \tilde{x}_0)e^{-\delta t}} \left( \frac{\delta}{\lambda - \delta} (e^{-\delta t} - e^{-\lambda t}) \dot{q}_0 + (1 - e^{-\delta t})\Delta \log(q) \right)
\]

Notice that Proposition 1 implies that

\[
\Delta \log(q) = \frac{1}{\chi} (\theta - 1) \Delta \log A
\]

\[
\Delta \log(q_0) = \frac{1}{\chi} \left( 1 + \frac{\lambda - \delta}{\delta} \right) (\theta - 1) \Delta \log A
\]

\[
\dot{q}_0 = \Delta \log(q_0) - \Delta \log(q) = \frac{1}{\chi} \frac{\lambda - \delta}{\delta} (\theta - 1) \Delta \log A
\]

Thus,

\[
\log \left( \frac{e^{\text{young}}_{0}}{e^{\text{old}}_{0}} \right) - \log \left( \frac{e^{\text{new}}_{0}}{e^{\text{old}}_{0}} \right) \approx \frac{\psi}{\chi} \frac{1}{1 - (1 - \tilde{x}_0)e^{-\delta t}} \left( e^{-\delta t} - e^{-\lambda t} + (1 - e^{-\delta t}) \right) (\theta - 1) \log A
\]

\[
\approx \frac{\psi}{\chi} \frac{1 - e^{-\lambda t}}{1 - (1 - \tilde{x}_0)e^{-\delta t}} (\theta - 1) \Delta \log A,
\]

which is equivalent to (2).

### 2.2 Parameterization by impulse response matching

We now describe how to parametrize our theory to match the empirical impulse response functions in Section 6. To this end, we map the $H$ technology in our theory to the set of production activities performed by cognitive-intensive occupations. We calibrate our theory in two steps. In the first step, we exogenously specify a subset of parameters and functions in the theory. We set the discount rate to match an annual interest rate of 2%, $\rho = 0.02$. We calibrate the elasticity of substitution across cognitive and non-cognitive intensive occupations to $\theta = 3$. Finally, for all welfare calculations, we specify welfare-weights $r e^{-rt}$ with $r = \rho + \delta$ so that the social discounting of future generations is identical to the discounting of worker’s future utility.

We also specify functional forms for the productivity of skill types in the two technologies. We abstract from differences in non-cognitive productivity across skills by normalizing $a(i) \equiv 1$. This implies that, for any given worker generation,
employment and payroll responses are both driven by the degree of technology-skill specificity in the economy.\footnote{The function form of \( a(i) \) controls how labor earnings respond to changes in the employment composition across technologies – for a discussion, see Adão (2016). Alternative specifications of \( a(i) \) can thus be used to match responses in relative earnings for different worker generations.} In addition, we assume that \( \sigma(i) \) takes the form of a logistic function:

\[
\sigma(i) = \frac{e^{\sigma(i-l)}}{1 + e^{\sigma(i-l)}}
\]

where \( l \) is the assignment threshold in the initial stationary equilibrium. This specification is a tractable manner of capturing technology-skill specificity in the economy. It implies that the equilibrium exists for any \( \sigma > 0 \) since the relative productivity is bounded. Also, by setting the midpoint of the function to \( l \), the parameter \( \sigma \) controls the elasticity of \( \sigma(i) \) for the marginal skill types in the initial equilibrium (i.e., \( i \) close to \( l \)). Thus, \( \sigma \) specifies the magnitude of technology-skill specificity, \( 1/\eta \).

In the second step, we use the estimated responses of Section 6 to calibrate \((\delta, \sigma, \psi)\). In doing so, we select the distribution of innate ability to normalize the initial skill distribution to be uniform: \( s_0(i) \equiv 1 \).\footnote{In this calibration, we select the distribution of innate ability distribution, \( s(i) \), to generate a uniform distribution of skills in the initial equilibrium: \( s_0(i) \equiv 1 \). In our theory, this normalization is innocuous since it does not affect changes in the skill distribution for a given change in \( q \) conditional on setting \( \eta \) to match the short-run employment change.} We formally present the parametrization procedure next, along with an analysis of the model fit. For all parameters, we assume that the shock starts with the roll-out of broadband internet in 2003. We then select parameters to match the estimates for the period of 2008 to 2014 in which we find statistically significant response in the relative payroll and relative employment of cognitive-intensive occupations.

**Generation size:** \( \delta \) and \( \bar{x}_0 \). We first set \( \bar{x}_0 \) to match the 60% share of young workers in the national population in 1997. We then select \( \delta \) to match the incline of 25 p.p. in the share of young workers in population between 1997 and 2014. Specifically, we select \( x \) and \( \delta \) such that

\[
\hat{\delta} = \frac{1}{2014 - 1997} \log(0.40/0.15)
\]

\[
x = -\frac{1}{\hat{\delta}} \log 0.4.
\]

We obtain \( \delta = 0.0574 \). This says that the expected work life of a worker after turning 40 years is 18 further years.

**Rate of convergence:** \( \lambda \). Proposition 1 implies that it is possible to write the

\( \)
impulse response function of relative output as

$$\Delta \log(y_t) = \alpha_0 + \alpha_1 e^{-\lambda t}$$

where \(\alpha_0 > 0, \alpha_1 < 0,\) and \(\lambda > 0.\)

We select the parameter \(\lambda\) to match the growth in the estimates response of relative payroll of more cognitive-intensive occupations:

$$\hat{\lambda} = \arg \min_{\lambda} \sum_{t=2008}^{2014} \left[ \left( \hat{\beta}^y_t - \hat{\beta}^y_{2007} \right) - \alpha_1 e^{-\lambda(t-2007)} \right]^2$$

(4)

where \(\hat{\beta}^y_t\) are the estimated coefficient reported in Panel B of Figure 6.

The minimization problem in (4) yields \(\hat{\lambda} = 0.135.\) Figure 1 shows the fit of the calibrated model

**Cost of skill investment: \(\psi.\)** Theorem 1 implies that

$$\kappa \eta = \psi \hat{\lambda} - \theta$$

(5)

where

$$\alpha = \delta \left[ \left( \frac{\rho}{2} + \lambda \right)^2 - \left( \frac{\rho}{2} \right)^2 - \delta (\rho + \delta) \right]^{-1}.$$ 

(6)

Using expression (2), we have that

$$\Delta \log e_i^{young} - \Delta \log e_i^{old} = \frac{\psi}{\chi} \frac{1 - e^{-\lambda t}}{1 - (1 - \tilde{x}_0) e^{-\delta t}} (\theta - 1) \Delta \log A.$$ 

From Proposition 1,

$$(\theta - 1) \Delta \log(A) = \Delta \log(y_t) \left( \frac{1 + \kappa \eta \chi}{\theta + \kappa \eta} + \frac{\psi \theta - 1}{\chi \theta + \kappa \eta} (1 - e^{-\lambda t}) \right)^{-1}$$

(7)

where \(\chi = (\theta + \kappa \eta) (\rho + \delta) + \psi.\)

Combining these two expressions, we get that

$$\frac{\Delta \log e_i^{young} - \Delta \log e_i^{old}}{\Delta \log y_t} = \frac{1 - e^{-\lambda t}}{1 - (1 - \tilde{x}_0) e^{-\delta t}} \left( \frac{1 + \kappa \eta \chi}{\theta + \kappa \eta} \frac{\psi \theta - 1}{\chi \theta + \kappa \eta} (1 - e^{-\lambda t}) \right)^{-1}.$$ 

(8)

Using the expression for \(\kappa \eta\) in (5),

$$\frac{\Delta \log e_i^{young} - \Delta \log e_i^{old}}{\Delta \log y_t} = \frac{1 - e^{-\lambda t}}{1 - (1 - \tilde{x}_0) e^{-\delta t}} \left( \rho + \delta \right) \left( \frac{1 + \psi \alpha - \theta}{\psi} + 1 - \frac{\theta - 1}{\psi \alpha} e^{-\lambda t} \right)^{-1}.$$ 

(9)
We then define the function:

\[ F_\psi(\psi, t) \equiv \frac{1 - e^{-\lambda t}}{1 - (1 - \hat{x}_0)e^{-\delta t}} \left( (\rho + \delta) \frac{1 + \psi \hat{\lambda} - \theta}{\psi} + 1 - \theta - 1 \right)^{-1}. \] (10)

To calibrate \( \psi \), we first our calibrated values of \((\lambda, \delta, \rho)\) to compute \( \alpha \) using (6). Our baseline calibration implies that \( \hat{\alpha} = 3.484 \). We then select the parameter \( \psi \) to match the ratio of the between-generation employment response and the payroll response:

\[ \hat{\psi} = \arg \min_{\psi} \sum_{t=2008}^{2014} \left[ \frac{\hat{\beta}^{\text{young}}_t - \hat{\beta}^{\text{old}}_t}{\hat{\beta}^{\text{young}}_t} - F_\psi(\psi, t) \right]^2 \] (11)

where \( \hat{\beta}^{\text{young}}_t \) are the estimated coefficients reported in Panel B of Figure 6, and \( \hat{\beta}^{\text{young}} - \hat{\beta}^{\text{old}} \) is the between-generation employment response obtained with the estimated coefficients reported in Panel A of Figure 6.

The minimization problem in (11) yields \( \hat{\psi} = 0.345 \). Figure 2 shows the fit of the calibrated model.

**Technology-skill specificity: \( \eta \).** The combination of (1) and (7) implies that

\[ \frac{\Delta \log e^{\text{old}}_t}{\Delta \log y_t} \approx \frac{\eta}{e_{H,0-1}} \frac{1 - \frac{\psi}{\hat{\lambda}}(1 - e^{-\lambda t})}{1 + \kappa \eta + \frac{\psi}{\hat{\lambda}}(\theta - 1)(1 - e^{-\lambda t})}. \]

Using the expression for \( \kappa \eta \) in (5),

\[ \frac{\Delta \log e^{\text{old}}_t}{\Delta \log y_t} \approx \frac{\eta}{e_{H,0-1}} \frac{1 - \frac{(\theta-1)(1 - e^{-\lambda t})}{\alpha + (\rho + \psi)(\hat{\lambda} + 1)}}{1 + \psi \hat{\alpha} - \theta + \frac{(\theta-1)(1 - e^{-\lambda t})}{\alpha + (\rho + \psi)(\hat{\lambda} + 1)}}. \] (12)

We then define

\[ F_\eta(\eta, t) \equiv \frac{\eta}{e_{H,0-1}} \frac{1 - \frac{\theta-1}{\hat{\lambda} + 1}(1 - e^{-\lambda t})}{1 + \hat{\psi} \hat{\alpha} - \theta + \frac{\theta-1}{\hat{\lambda} + 1}(1 - e^{-\lambda t})}. \] (13)

where \((\hat{\delta}, \hat{\lambda}, \hat{\psi})\) are the calibrated parameters above and \( e_{H,0-1} \) is the initial share of employment in cognitive-intensive occupations.

We select the parameter \( \eta \) to match the ratio of the employment response of old workers and the payroll response:

\[ \hat{\eta} = \arg \min_{\eta} \sum_{t=2008}^{2014} \left[ \frac{\hat{\beta}^{\text{old}}_t}{\hat{\beta}^{\text{young}}_t} - F_\eta(\eta, t) \right]^2. \] (14)
\( \hat{\beta}_y \) are the estimated coefficients reported in Panel B of Figure 6, and \( \hat{\beta}_t^{old} \) are the estimated coefficients reported in Panel A of Figure 6.

The negative point estimates reported in Panel A of Figure 6 imply that the minimization problem in (14) yields \( \hat{\eta} < 0 \). Since the employment response of old generations is small and nonsignificant, we assume that they are identical to zero, which yields \( \hat{\eta} = 0 \). Hence, we calibrate \( \eta = 0.01 \) and evaluate the model predictions under alternative specifications of this parameter.

![Figure 1. Calibration of \( \lambda \)](image)

Note. Blue dots represent the point estimates reported in Panel B of Figure 6. Black solid curve represents the best fit line with \( \lambda = 0.135 \) obtained from the solution of (4).
2.3 Dynamic responses to cognitive-biased technological innovations

We now present the dynamic responses in our parametrized model. As in Section 7, we evaluate a shock to $A$ that leads to an increase in the employment share in cognitive-intensive occupations from 20% to 50%. Figure 3 presents the results.

Consider first the response at $t = 0$. Given that our theory abstracts from several additional sources of dynamics, it would be wrong to interpret the impact adjustment as happening instantaneously in reality. We view this short-run response as capturing changes over the time window encompassing dynamic forces triggered by other variables that are likely to move faster than the distribution of skills (e.g., physical capital). In other words, we prefer to interpret the “length” of the impact adjustment as related to the time that it takes for such faster moving variables to converge to the new long-run equilibrium. Results show that there is a substantial increase in the relative cognitive-intensive output in the short-run. This large response is a consequence of the large magnitude of the shock. This becomes clear when we take into account that relative employment almost does not change at impact because of the high technology-skill specificity (i.e., $\eta \approx 0$). The combination of the large increase in relative output and the small increase in relative employment translates into large changes in lifetime inequality.
Our results also indicate that the responses in all outcomes change substantially over time (measured in terms of worker generations, \(1/\delta \approx 18\) yrs). Over the course of the two generations following the shock, the responses in relative output doubles in magnitude due to the reallocation of workers across technologies. Such a reallocation is entirely driven by incoming generations of young workers. This pattern is a consequence of the change in the skill distribution across generations. The bottom right panel shows that the initial spike in lifetime inequality induces young workers to invest in high-\(i\) skills allocated to cognitive-intensive occupations. This gives rise to substantial skill heterogeneity across generations. As young generations replace old generations, the economy’s skill distribution becomes more biased towards high-\(i\) types, leading to a large decline in the present value of the relative wage in cognitive-intensive occupations (which recedes by more than 30% over the course of two generations).

Figure 3. Dynamic responses to a cognitive-biased innovation at \(t = 0\)

Note. The figure reports the theoretical impulse response function with a shock calibrated to increase the employment share in cognitive-intensive occupations from 20% to 50% between stationary equilibria. Baseline calibration described in Appendix B.5.
3. Additional Results

3.1 Microfoundation of the Production Functions in (4)–(5)

Consider two firms: high-tech \((k = H)\) and low-tech \((k = L)\). Assume that the output of firm \(k\) at time \(t\) aggregates per-worker output \(x_{kt}(i)\),

\[
X_{kt} = \int_0^1 x_{kt}(i)s_{kt}(i)di,
\]

where \(s_{kt}(i)\) is the quantity demanded of workers of type \(i\) at time \(t\) by firm \(k\).

The output of workers of type \(i\) depends on their skills to perform cognitive and noncognitive tasks, \(\{a_C(i), a_{NC}(i)\}\), as well as how intensely each task is used in the firm’s production process:

\[
x_{kt}(i) = a_C(i)\beta_k a_{NC}(i)^{1-\beta_k},
\]

where \(\beta_k\) denotes the production intensity of firm \(k\) on cognitive tasks.

In our model, technology-skill specificity arises whenever firms are heterogeneous in terms of task intensity and workers are heterogeneous in terms of their task bundle. To see this, suppose that firm \(H\)’s technology uses cognitive tasks more intensely than firm \(L\)’s technology, \(\beta_H > \beta_L\), and that a worker of type \(i\) is able to produce a higher cognitive-noncognitive task ratio than a worker of type \(j\), \(a_C(i)/a_{NC}(i) > a_C(j)/a_{NC}(j)\). In this case, \(i\) has a higher relative output with the cognitive-intensive technology \(H\) than \(j\), \(x_{Ht}(i)/x_{Lt}(i) > x_{Ht}(j)/x_{Lt}(j)\), and, therefore, type \(i\) is more complementary to the cognitive-intensive technology \(H\) than type \(j\).

To map this setting to the production functions in (4)–(5), we assume that high-tech production is more intensive in cognitive tasks than low-tech production, \(\beta_H > \beta_L\). We also assume that types differ in terms of their skill bundle and, without loss of generality, impose that high-\(i\) types are relatively better in performing cognitive-intensive tasks.

1. High-tech technology \(H\) uses cognitive tasks more intensely than Low-tech technology \(L\): \(\beta_H > \beta_L\).

2. Define \(\sigma(i) \equiv \left(\frac{a_C(i)}{a_{NC}(i)}\right)^{\beta_H-\beta_L}\) and \(\alpha(i) \equiv a_C(i)^{\beta_L}a_{NC}(i)^{1-\beta_L}\). Assume that high-\(i\) types have higher cognitive-noncognitive task ratio: \(\sigma(i)\) is increasing in \(i\).

3.2 Welfare Consequences of Adjustment Across Generations

This section investigates how calculations of the welfare consequences of technological shocks are affected by the speed of adjustment of labor market outcomes.
along the transition to the new equilibrium. In our theory, the transitional dynamics arise from the changes in the skill distribution. So, in order to evaluate its consequences, we consider a static version of our model in which we shut down any skill investment of young workers. However, we allow this static model to match labor market responses over one particular horizon. This exercise thus speaks directly to the risks of ignoring the adjustment across generations by focusing on estimates of the impact of new technologies on labor market outcomes over fixed time horizons.

To be more precise, we engage in the following thought experiment. Consider an economy subject to a one-time permanent shock $\Delta \log A$. Suppose that this economy behaves according to the theoretical predictions described in Section 3 with short- and long-run skill supply elasticity given by $\eta$ and $\psi$, respectively. We consider a researcher that relies on a static assignment model to analyze how this economy responds to the technological shock. Through the lens of our theory, this researcher considers a misspecified parametrization of the economy in which the long-run elasticity equals zero. This parametrization shuts down any dynamics in the economy because the skill distribution is the same for all generations. We assume that this researcher observes responses in labor market outcomes over a fixed horizon $t = T$. We focus on changes in lifetime inequality since this is the main endogenous outcome entering the welfare computations in Proposition 2. We consider two ways in which the researcher may decide to use the static model to match the observed inequality response, $\Delta \log q_T$. In the first approach, the researcher observes the true shock ($\Delta \log A^1 = \Delta \log A$), and selects $\eta^1$ to match $\Delta \log q_T$ with $\psi^1 = 0$. In the second, the researcher observes the true parameter ($\eta^2 = \eta$), and selects the size of the shock $\Delta \log A^2$ to match $\Delta \log q_T$ with $\psi^2 = 0$.

The following proposition shows that, despite matching inequality responses at time $T$, this researcher misses the economy’s transitional dynamics triggered by the evolution of the skill distribution across generations. This introduces biases in the evaluation of the welfare consequences of the technological innovation.

In particular, it shows that there are multiple ways in which researchers can use a static version of our model to match observed inequality responses over a fixed horizon. All versions ignore the transitional dynamics of labor market outcomes generated by changes in the skill distribution across generations. This introduces biases in the evaluation of the welfare consequences of new technologies. If the researcher only matches inequality responses in short horizons (i.e. $T$ is low), then she will think that inequality will remain high in the future. This makes her overpredict the present value of lifetime inequality, and underpredict the average welfare gain. Alternatively, a researcher using the first approach would reach the opposite conclusions if she matches inequality responses in long horizons (i.e. $T$ is high).

Such biases will be larger when the adjustment is slower due to the larger changes in the skill distribution along the transition. As shown in Section 4, this
is the case whenever the skill supply elasticity is low in the short-run (i.e, $\eta$ is low) but large in the long-run (i.e., $\psi$ is large).

**Proposition 4** Consider an economy in which $\eta$ and $\psi$ are positive. Assume that $\Delta \log A$ generates a change in lifetime inequality between $t = 0$ and $t = T$ of $\Delta \log q_T$. Consider predictions under two alternative static parametrizations of the model ($\psi^1 = \psi^2 = 0$).

1. Suppose $\Delta \log A^1 = \Delta \log A$ is known such that $\frac{\Delta \log(A)}{\Delta \log(q_T)} > \frac{\theta(\rho + \delta)}{\theta - 1}$. There exists $\eta^1$ that matches $\Delta \log q_T$ with an associated $T^1$ such that $\Delta \bar{\Omega}^1 > \Delta \bar{\Omega}$ and $\Delta \bar{U}^1 < \Delta \bar{U}$ if, and only if, $T < T^1$.

2. Suppose $\eta^2 = \eta$ is known. There exists $\Delta \log A^2$ that matches $\Delta \log q_T$ with an associated $T^2$ such that $\Delta \bar{\Omega}^2 > \Delta \bar{\Omega}$ and $\Delta \bar{U}^2 < \Delta \bar{U}$ if $T < T^2$.

**Proof.**

We start by pointing out that, by the definition in Theorem 1, $\lambda^1 = \lambda^2 = \delta$ because $\psi^1 = \psi^2 = 0$. Thus, Proposition 1 immediately implies that both parametrizations must satisfy the condition that

$$\Delta \log(q_T) = \frac{\theta - 1}{(\theta + \kappa \eta^P)(\rho + \delta)} \Delta \log A^p$$  \hspace{1cm} (B.2)

where $P = 1$ for the first approach or $P = 2$ for the second approach.

Notice also that the combination of Theorem 1 and Proposition 2 implies that

$$\Delta \bar{\Omega} = (\rho + \delta) \Delta \log(q_T) - (\rho + \delta) \hat{q}_0 \left( e^{-\lambda T} - 1 + \frac{\lambda}{r + \lambda} \right)$$

and, therefore,

$$\Delta \bar{\Omega} = (\rho + \delta) \Delta \log(q_T) \left( \frac{1 + \frac{\lambda - \delta}{\delta} \frac{r}{r + \lambda}}{1 + \frac{\lambda - \delta}{\delta} e^{-\lambda T}} \right).$$  \hspace{1cm} (B.3)

This expression implies that, because $\lambda^1 = \lambda^2 = \delta$, both parametrizations entail

$$\Delta \bar{\Omega}^P = (\rho + \delta) \Delta \log(q_T)$$  \hspace{1cm} (B.4)

for $P = 1, 2$.

We now use these expressions two establish the two parts of the proposition.

**Part 1.** In the first approach, we set $\Delta \log A^1 = \Delta \log A$. So, by equation (B.2), we must set

$$\kappa \eta^1 = \frac{\theta - 1}{\rho + \delta} \Delta \log(A) - \theta,$$

which is positive as long as $\frac{\Delta \log(A)}{\Delta \log(q_T)} > \frac{\theta(\rho + \delta)}{\theta - 1}$.
By taking the ratio between the expressions in (B.3) and (B.4),
\[
\frac{\Delta \bar{\Omega}_1}{\Delta \bar{\Omega}} > 1 \iff e^{-\lambda T} > \frac{r}{r + \lambda} \iff T < T^1 \equiv \frac{1}{\lambda} \log \left( \frac{r + \lambda}{r} \right).
\]

The expression of \( \Delta \bar{U} \) in Proposition 2 immediately implies that \( \Delta \bar{\Omega}_1 > \Delta \bar{\Omega} \iff \Delta \bar{U}_1 < \Delta \bar{U} \) whenever \( y_\infty > e_\infty \).

**Part 2.** In the second approach, we set \( \eta^2 = \eta \). So, by equation (B.2), we must set
\[
\Delta \log A^2 = \Delta \log (q_T) \frac{(\theta + \kappa \eta)(\rho + \delta)}{\theta - 1}.
\]

Expressions in (B.3) and (B.4) also hold in this case, so the same steps used above guarantee that \( \Delta \bar{\Omega}_2 > \Delta \bar{\Omega} \) if, and only if, \( T < T^1 \). To establish the result, it is sufficient to show that \( \Delta \log A^2 \leq \Delta \log A \) because, by Proposition 2, \( \Delta \bar{\Omega}^2 > \Delta \bar{\Omega} \) and \( \Delta \log A^2 \leq \Delta \log A \) imply that \( \Delta \bar{U}_2 < \Delta \bar{U} \).

We now show that \( \Delta \log A^2 \leq \Delta \log A \). By combining Proposition 1 and equation (B.2), we have that
\[
\Delta \log A^2 = \frac{(\theta + \kappa \eta)}{(\theta + \kappa \eta + \frac{\psi}{\rho + \delta})} \left( 1 + \frac{\lambda - \delta}{\delta} e^{-\lambda T} \right) \Delta \log A
\]
and, therefore,
\[
\Delta \log A^2 \leq \frac{(\theta + \kappa \eta)}{(\theta + \kappa \eta + \frac{\psi}{\rho + \delta})} \frac{\lambda}{\delta} \Delta \log A.
\]

So, \( \Delta \log A^2 \leq \Delta \log A \) if
\[
F(\psi) \equiv \frac{(\theta + \kappa \eta)}{(\theta + \kappa \eta + \frac{\psi}{\rho + \delta})} \frac{\lambda(\psi)}{\delta} \leq 1
\]
with \( \lambda(\psi) \) defined in Theorem 1.

This condition always holds because \( \lambda(0) = \delta, F(0) = 1 \) and \( \text{sign} \left( \frac{\partial F(\psi)}{\partial \psi} \right) < 0 \).
To see this, we use the expression for $\lambda(\psi)$ in Theorem 1 to show that

\[
sign \left( \frac{\partial F(\psi)}{\partial \psi} \right) = \text{sign} \left( \frac{\partial \lambda(\psi)}{\partial \psi} \left( \theta + \kappa \eta + \frac{\psi}{\rho + \delta} \right) - \frac{\lambda}{\rho + \delta} \right)
\]
\[
= \text{sign} \left( \frac{1}{2\lambda + \rho} \frac{\delta}{\theta + \kappa \eta} \left( \theta + \kappa \eta + \frac{\psi}{\rho + \delta} \right) - \frac{\lambda}{\rho + \delta} \right)
\]
\[
= \text{sign} \left( \frac{1}{2\lambda + \rho} \left( \delta + \frac{\delta}{\rho + \delta} \left( \theta + \kappa \eta \right) \right) - \frac{\lambda}{\rho + \delta} \right)
\]
\[
= \text{sign} \left( \frac{1}{2\lambda + \rho} \left( \frac{\delta}{\rho + \delta} \left[ \left( \lambda + \rho \right)^2 - \left( \frac{\rho}{2} \right)^2 - \delta(\rho + \delta) \right] \right) - \frac{\lambda}{\rho + \delta} \right)
\]
\[
= \text{sign} \left( \frac{1}{2\lambda + \rho} \left[ \left( \lambda + \rho \right)^2 - \left( \frac{\rho}{2} \right)^2 \right] - \lambda \right)
\]
\[
= \text{sign} \left( \frac{\lambda + \rho}{2\lambda + \rho} - 1 \right).
\]