Supplemental Appendix
Self-Selection and the Diminishing Returns of Research

By Lorenz K.F. Ekerdt and Kai-Jie Wu

Last updated: July 2026

A Model appendix

A.1 Full semi-endogenous growth model

This section describes the demand-side of the semi-endogenous growth model described in Section II, and relates it to a textbook semi-endogenous growth model without selection into research (Acemoglu, 2009). To allow for decentralization of the model, first suppose that the stock of ideas A_t enables the production of a continuum of intermediate input varieties with measure equal to A_t. Entry into research is free and discovering an idea enables an entrant to obtain a perpetual patent to the intermediate input variety corresponding to that idea. In the final goods sector, a representative firm uses low-skilled production labor, high-skilled production labor, and the continuum of intermediate input varieties to produce the final consumption good; its profit-maximization problem is thus

$$\max_{x_{it},Z_{Ht},Z_{Nt}} \frac{1}{1-\alpha-\beta}\int_0^{A_t} x_{it}^{1-\alpha-\beta} di Z_{Ht}^{\beta} Z_{Nt}^{\alpha} di - W_{Ht}Z_{Ht} - W_{Nt}Z_{Nt} - \int_0^{A_t} p_{it}x_{it}di. \tag{A1}$$

The aggregate production function thus takes a simple Cobb-Douglas form; this assumption is inessential but simplifies the characterization of the model significantly. Conditional on entry, the intermediate input monopolists can produce their variety at constant marginal cost ψ. The consumer side is standard: a representative household with discount rate ρ and instantaneous utility function $\frac{C_t^{1-\sigma}-1}{1-\sigma}$ chooses consumption and savings given their budget constraint and real interest rate r_t. Let $\mathcal{R}_t = \frac{W_{Rt}}{W_{Ht}}, \mathcal{R}_t^N = \frac{W_{Rt}}{W_{Nt}}$, then we can summarize the model with the following set of equations:

$$\text{Idea production} \quad \dot{A}_t = \chi A_t^{\phi} Z_{Rt}^{\lambda} \tag{A2}$$

$$\text{Researcher labor supply} \quad Z_{Rt}^S = L_t\Gamma_R\left(\mathcal{R}_t, \mathcal{R}_t^N\right) \tag{A3}$$

$$\text{High-skill production labor supply} \quad Z_{Ht}^S = L_t\Gamma_H\left(\mathcal{R}_t, \mathcal{R}_t^N\right) \tag{A4}$$

$$\text{Low-skill production labor supply} \quad Z_{Nt}^S = L_t\Gamma_N\left(\mathcal{R}_t, \mathcal{R}_t^N\right) \tag{A5}$$

$$\text{Optimal intermediate input price} \quad p_{it} = \frac{\psi}{1-\alpha-\beta} \tag{A6}$$

$$\text{Euler equation} \quad \frac{\dot{c}_t}{c_t} = \frac{1}{\sigma}(r_t - \rho) \tag{A7}$$

$$\text{High-skill production labor demand} \quad Z^D_{Ht} = \left[\frac{\beta Z^{\alpha}_{Nt}\tilde{A}_t}{1-\alpha-\beta}\right]^{\frac{1}{1-\beta}} W_{Ht}^{-\frac{1}{1-\beta}} \tag{A8}$$

$$\text{Low-skill production labor demand} \quad Z^D_{Nt} = \left[\frac{\alpha Z^{\beta}_{Ht}\tilde{A}_t}{1-\alpha-\beta}\right]^{\frac{1}{1-\alpha}} W_{Nt}^{-\frac{1}{1-\alpha}} \tag{A9}$$

$$\text{Intermediate input demand} \quad x^D_{it} = Z_{Nt}^{\frac{\alpha}{\alpha+\beta}} Z_{Ht}^{\frac{\beta}{\alpha+\beta}} p_{it}^{-\frac{1}{\alpha+\beta}} \tag{A10}$$

$$\text{Free entry into research} \quad W_{Rt} = \chi\lambda A_t^{\phi} Z_{R,t}^{\lambda-1} V_{it} \tag{A11}$$

with c_t per capita consumption, V_{it} the value of the perpetual patent for intermediate input i at time t, $\tilde{A}_t \equiv \int_0^{A_t} x_{it}^{1-\alpha-\beta} di$, and $\Gamma_j(\cdot)$ denoting the function that maps the joint distribution of (z_{Ci}, z_{Ri}, z_{Hi}) to labor productivity in sector j. An equilibrium is then a set of allocations $\{Z_{Nt}, Z_{Rt}, Z_{Nt}, C_t\}$, a time path of ideas A_t, and a set of prices $\{p_{it}, r_t, W_{Rt}, W_{Ht}, W_{Nt}\}$ that satisfy these equations.

A.2 Solving the balanced growth path (BGP)

Normalizing $p_{it} = 1$, the instantaneous profit function is

$$\pi_t = Z_{Ht}^{\frac{\beta}{\alpha+\beta}} Z_{Nt}^{\frac{\alpha}{\alpha+\beta}} (1-\psi). \tag{A12}$$

We can also solve

$$\tilde{A}_t = Z_{Nt}^{\frac{\alpha(1-\alpha-\beta)}{\alpha+\beta}} Z_{Ht}^{\frac{\beta(1-\alpha-\beta)}{\alpha+\beta}} A_t \tag{A13}$$

Define a BGP to be a path of allocations under which consumption and output grow at the same constant rate, and denote with $*$ the values of equilibrium objects along this BGP. Suppose, for now, that there is no utility costs that changes the sector choices, i.e., $\mu_{Ht} - \mu_{Rt} = \mu_{Rt} + \mu_{C,t} - \mu_{Nt} = 0$; once we have solved the BGP under this assumption, relaxing it is straightforward, see equation (A29). On a BGP, the Hamilton-Jacobi-Bellman equation must satisfy

$$(r^* - g_L)V^*_{it} = \pi^*_t. \tag{A14}$$

Combining these equations with the labor demand equations stemming from the final goods producer's profit maximization problem yields:

$$W_{Rt}^{*} = \frac{\lambda\chi(1-\psi)}{r^{*}-g_L} A_t^{*\phi} Z_{Rt}^{*\lambda-1} Z_{Ht}^{*\frac{\beta}{\alpha+\beta}} Z_{Nt}^{*\frac{\alpha}{\alpha+\beta}} \tag{A15}$$

$$W_{Ht}^{*} = \frac{\beta}{1-\alpha-\beta} Z_{Nt}^{*\frac{\alpha}{\alpha+\beta}} Z_{Ht}^{*-\frac{\alpha}{\alpha+\beta}} A_t^{*} \tag{A16}$$

$$W_{Nt}^{*} = \frac{\alpha}{1-\alpha-\beta} Z_{Nt}^{*-\frac{\beta}{\alpha+\beta}} Z_{Ht}^{*\frac{\beta}{\alpha+\beta}}. \tag{A17}$$

Along a balanced growth path, the employment shares of the different types of labor must be constant; we therefore have

$$\mathcal{R}^{*} = \frac{W_{Rt}^{*}\exp(\mu_{Rt})}{W_{Ht}^{*}\exp(\mu_{Ht})} = \frac{\chi\lambda(1-\alpha-\beta)(1-\psi)}{\beta(r^{*}-g_L)} A_t^{*\phi-1} Z_{Rt}^{*\lambda-1} Z_{Ht}^{*} \tag{A18}$$

$$\mathcal{R}_N^{*} = \frac{W_{Rt}^{*}\exp(\mu_{Rt})}{W_{Nt}^{*}\exp(\mu_{Nt}-\mu_{Ct})} = \frac{\chi\lambda(1-\alpha-\beta)(1-\psi)}{\alpha(r^{*}-g_L)} A_t^{*\phi-1} Z_{Rt}^{*\lambda-1} Z_{Nt}^{*}. \tag{A19}$$

Rearranging the idea production function yields $A_t^{*\phi-1} = g_A^{*}\left(\chi Z_{Rt}^{*\lambda}\right)^{-1}$ and thus we can rewrite the above as:

$$\mathcal{R}^{*} = \frac{\lambda(1-\alpha-\beta)(1-\psi)}{\beta(r^{*}-g_L)} \frac{Z_{Ht}^{*}}{Z_{Rt}^{*}} \tag{A20}$$

$$\mathcal{R}_N^{*} = \frac{\lambda(1-\alpha-\beta)(1-\psi)}{\alpha(r^{*}-g_L)} \frac{Z_{Nt}^{*}}{Z_{Rt}^{*}}. \tag{A21}$$

Finally, plugging in the labor supply equations:

$$\mathcal{R}^{*} = \frac{(1-\alpha-\beta)(1-\psi)}{\beta(r^{*}-g_L)} \frac{\Gamma_H(\mathcal{R}^{*},\mathcal{R}_N^{*})}{\Gamma_R(\mathcal{R}^{*},\mathcal{R}_N^{*})} \tag{A22}$$

$$\mathcal{R}_N^{*} = \frac{(1-\alpha-\beta)(1-\psi)}{\alpha(r^{*}-g_L)} \frac{\Gamma_N(\mathcal{R}^{*},\mathcal{R}_N^{*})}{\Gamma_R(\mathcal{R}^{*},\mathcal{R}_N^{*})}. \tag{A23}$$

To solve for r^{*}, note that we can rewrite the idea production function as

$$\frac{\dot{A}_t^{*}}{A_t^{*}} = \chi A_t^{*\phi-1} Z_{Rt}^{*\lambda} \tag{A24}$$

Taking logs, deriving with respect to time, and rearranging yields

$$g_A^{*} = \frac{\lambda g_{Z_R}^{*}}{1-\phi} = \frac{\lambda g_L}{1-\phi} \tag{A25}$$

Plugging this into the Euler equation gives

$$r^* = \frac{\lambda\sigma}{1-\phi} g_L + \rho, \tag{A26}$$

and so equations (A22) and (A23) can be rewritten

$$\mathcal{R}^* = \frac{\lambda(1-\alpha-\beta)(1-\psi)}{\beta(\frac{\lambda\sigma-(1-\phi)}{\phi} g_L + \rho)} \frac{\Gamma_H(\mathcal{R}^*, \mathcal{R}_N^*)}{\Gamma_R(\mathcal{R}^*, \mathcal{R}_N^*)} \tag{A27}$$

$$\mathcal{R}_N^* = \frac{\lambda(1-\alpha-\beta)(1-\psi)}{\alpha(\frac{\lambda\sigma-(1-\phi)}{\phi} g_L + \rho)} \frac{\Gamma_N(\mathcal{R}^*, \mathcal{R}_N^*)}{\Gamma_R(\mathcal{R}^*, \mathcal{R}_N^*)}. \tag{A28}$$

This yields two equations in two unknowns, which can be solved for $(\mathcal{R}^*, \mathcal{R}_C^*)$. Given this solution, the BGP allocations of all equilibrium objects can be found by unwinding the set of equation substitutions leading to equations (A27) and (A28). Note that this model has the exact same BGP growth rate as one without selection into research, thus justifying our description of selection effects as transitional. This can easily be verified by setting $z_{Ri} = z_P = z_{Ci} = 1$ for all workers and $\mu_{C,t} = 0$, in which case there is no self-selection but equation (A25) still holds.

Now we relax the assumption that $\mu_{Ht} - \mu_{Rt} = \mu_{Rt} + \mu_{C,t} - \mu_{Nt} = 0$. Let $**$ denote equilibrium values in the resulting BGP. Then if

$$\begin{aligned} \frac{W_{Rt}^{**}}{W_{Ht}^{**}} &= \frac{W_{Rt}^{*}}{W_{Ht}^{*}} \times \exp(\mu_{Rt} - \mu_{Ht}) \\ \frac{W_{Rt}^{**}}{W_{Nt}^{**}} &= \frac{W_{Rt}^{*}}{W_{Nt}^{*}} \times \exp(\mu_{Rt} + \mu_{C,t} - \mu_{Nt}), \end{aligned} \tag{A29}$$

We have $\mathcal{R}^{**} = \mathcal{R}^*, \mathcal{R}_N^{**} = \mathcal{R}_N^*$, and so, with the exception of the relative wage rates, the BGP is unchanged.[A1]

A.3 Evolution of average ability

To derive equation (10) in the main text, first write

$$\bar{z}_{Rt} = \mathbb{E}\left[z_{Ci} + z_{Ri} \mid z_{C,t}^m \leq z_{Ci}, z_{R,t}^m(z_{Hi}) \leq z_{Ri}\right] = \frac{z_{Rt}}{S_{Rt}} \tag{A30}$$

with

$$z_{Rt} = \int_{-\infty}^{\infty} \int_{z_{Rt}^m(z_{Hi})}^{\infty} \int_{z_{Ct}^m}^{\infty} (z_{Ri} + z_{Ci}) f(z_{Ri}, z_{Hi}, z_{Ci}) dz_{Ci} dz_{Ri} dz_{Hi} \tag{A31}$$

[A1]Clearly, this BGP will break down if either $\exp(\mu_{Rt} - \mu_{Ht}) \to \infty$ or $\exp(\mu_{Rt} + \mu_{C,t} - \mu_{Nt}) \to \infty$.

$$S_{Rt} = \int_{-\infty}^{\infty} \int_{z_{Rt}^m(z_{Hi})}^{\infty} \int_{z_{Ct}^m}^{\infty} f(z_{Ri}, z_{Hi}, z_{Ci}) dz_{Ci} dz_{Ri} dz_{Hi}. \tag{A32}$$

Taking time derivatives using Leibniz's rule, we have

$$\begin{aligned}\dot{z}_{Rt} = &-\dot{z}_{Rt}^m \int_{-\infty}^{\infty} \int_{z_{Ct}^m}^{\infty} (z_{Rt}^m(z_{Hi}) + z_{Ci}) f(z_{Rt}^m(z_{Hi}), z_{Hi}, z_{Ci}) dz_{Ci} dz_{Hi} \\ &- \dot{z}_{Ct}^m \int_{-\infty}^{\infty} \int_{z_{Rt}^m(z_{Hi})}^{\infty} (z_{Ri} + z_{Ct}^m) f(z_{Ri}, z_{Hi}, z_{Ct}^m) dz_{Ri} dz_{Hi},\end{aligned} \tag{A33}$$

and

$$\begin{aligned}\dot{S}_{Rt} = &-\dot{z}_{Rt}^m \int_{-\infty}^{\infty} \int_{z_{Ct}^m}^{\infty} f(z_{Rt}^m(z_{Hi}), z_{Hi}, z_{Ci}) dz_{Ci} dz_{Hi} \\ &- \dot{z}_{Ct}^m \int_{-\infty}^{\infty} \int_{z_{Rt}^m(z_{Hi})}^{\infty} f(z_{Ri}, z_{Hi}, z_{Ct}^m) dz_{Ri} dz_{Hi}.\end{aligned} \tag{A34}$$

We can write

$$\begin{aligned}\frac{\dot{z}_{Rt}}{S_{Rt}} = &-\dot{z}_{Rt}^m \frac{\int_{-\infty}^{\infty} \int_{z_{Ct}^m}^{\infty} (z_{Rt}^m(z_{Hi}) + z_{Ci}) f(z_{Rt}^m(z_{Hi}), z_{Hi}, z_{Ci}) dz_{Ci} dz_{Hi}}{M(z_{Ri} = z_{Rt}^m(z_{Hi}), z_{Ci} \geq z_{Ct}^m)} \times \frac{M(z_{Ri} = z_{Rt}^m(z_{Hi}), z_{Ci} \geq z_{Ct}^m)}{S_{Rt}} \\ &- \dot{z}_{Ct}^m \frac{\int_{-\infty}^{\infty} \int_{z_{Rt}^m(z_{Hi})}^{\infty} (z_{Ri} + z_{Ct}^m) f(z_{Ri}, z_{Hi}, z_{Ct}^m) dz_{Ri} dz_{Hi}}{M(z_{Ri} \geq z_{Rt}^m(z_{Hi}), z_{Ci} = z_{Ct}^m)} \times \frac{M(z_{Ri} \geq z_{Rt}^m(z_{Hi}), z_{Ci} = z_{Ct}^m)}{S_{Rt}} \\ = &-\dot{z}_{Rt}^m \mathbb{E}\left[z_{Rt}^m(z_{Hi}) + z_{Ci} \mid z_{Ri} = z_{Rt}^m(z_{Hi}), z_{Ci} \geq z_{Ct}^m\right] \times \frac{M(z_{Ri} = z_{Rt}^m(z_{Hi}), z_{Ci} \geq z_{Ct}^m)}{S_{Rt}} \\ &- \dot{z}_{Ct}^m \mathbb{E}\left[z_{Ri} + z_{Ct}^m \mid z_{Ri} \geq z_{Rt}^m(z_{Hi}), z_{Ci} = z_{Ct}^m\right] \times \frac{M(z_{Ri} \geq z_{Rt}^m(z_{Hi}), z_{Ci} = z_{Ct}^m)}{S_{Rt}},\end{aligned} \tag{A35}$$

and

$$\frac{\dot{S}_{Rt}}{S_{Rt}} = -\dot{z}_{Rt}^m \frac{M(z_{Ri} = z_{Rt}^m(z_{Hi}), z_{Ci} \geq z_{Ct}^m)}{S_{Rt}} - \dot{z}_{Ct}^m \frac{M(z_{Ri} \geq z_{Rt}^m(z_{Hi}), z_{Ci} = z_{Ct}^m)}{S_{Rt}}, \tag{A36}$$

where M denotes probability mass. Combining,

$$\begin{aligned}\dot{\bar{z}}_{Rt} = &\frac{\dot{z}_{Rt}}{S_{Rt}} - \frac{\dot{S}_{Rt}}{S_{Rt}} \bar{z}_t^R \\ = &\dot{z}_{Rt}^m \frac{M(z_{Ri} = z_{Rt}^m(z_{Hi}), z_{Ci} \geq z_{Ct}^m)}{S_{Rt}} \left[\bar{z}_t^R - \mathbb{E}\left[z_{Ri} + z_{Ct}^m \mid z_{Ri} = z_{Rt}^m(z_{Hi}), z_{Ci} \geq z_{Ct}^m\right]\right] \\ &+ \dot{z}_{Ct}^m \frac{M(z_{Ri} \geq z_{Rt}^m(z_{Hi}), z_{Ci} = z_{Ct}^m)}{S_{Rt}} \left[\bar{z}_t^R - \mathbb{E}\left[z_{Ri} + z_{Ct}^m \mid z_{Ri} \geq z_{Rt}^m(z_{Hi}), z_{Ci} = z_{Ct}^m\right]\right].\end{aligned} \tag{A37}$$

Imposing $(z_{Ri}, z_{Hi}) \perp z_{Ci}$, we can simplify this to

$$\begin{aligned}\dot{\bar{z}}_{Rt} &= \dot{z}^m_{Rt}\frac{M(z_{Ri} = z^m_{Rt}(z_{Hi}))}{M(z_{Ri} \geq z^m_{Rt}(z_{Hi}))}\left[\mathbb{E}\left(z_{Ri} \mid z_{Ri} \geq z^m_{Rt}(z_{Hi})\right) - \mathbb{E}\left(z^m_{Rt}(z_{Hi})\right)\right] \\ &+ \dot{z}^m_{Ct}\frac{M(z_{Ci} = z^m_{Ct})}{M(z_{Ci} \geq z_{Ct})}\left[\mathbb{E}\left(z_{Ci} \mid z_{Ci} \geq z^m_{Ct}\right) - z^m_{Ct}\right].\end{aligned} \tag{A38}$$

To derive $\dot{z}^m_{Ct}$, recall that $z^m_{Ct} = \mu_{Ct} - u_{Cit}$ with $u_{Cit} = \mathbb{E}(\max_{j\in\{R,H\}}\{z_j + \tilde{w}_j\})$. We can write

$$\begin{aligned}u_{Cit} &= \int_{-\infty}^{\infty} f(z_{Hi}) \int_{-\infty}^{\infty} \max\{z_{Ri} + \tilde{w}_{Rt}, z_{Hi} + \tilde{w}_{Ht}\} f(z_{Ri}|z_{Hi}) dz_{Ri} dz_{Hi} \\ &= \int_{-\infty}^{\infty} f(z_{Hi})\Big[\int_{z_{Hi}+\tilde{w}_{Ht}-\tilde{w}_{Rt}}^{\infty} (z_{Ri} + \tilde{w}_{Rt}) f(z_{Ri}|z_{Hi}) dz_{Ri} \\ &+ \int_{-\infty}^{z_{Hi}+\tilde{w}_{Ht}-\tilde{w}_{Rt}} (z_{Hi} + \tilde{w}_{Ht}) dz_{Ri}\Big] dz_{Hi}\end{aligned} \tag{A39}$$

Taking a time derivative yields

$$\begin{aligned}\dot{u}_{Cit} &= \int_{-\infty}^{\infty} f(z_{Hi})\Big[(\dot{\tilde{w}}_{Rt} - \dot{\tilde{w}}_{Ht})(z_{Hi} + \tilde{w}_{Ht}) f(z^m_{Rt}(z_{Hi})|z_{Hi}) + \int_{z^m_{Rt}(z_{Hi})}^{\infty} \dot{\tilde{w}}_{Rt} f(z_{Ri}|z_{Hi}) dz_{Ri} \\ &+ (\dot{\tilde{w}}_{Ht} - \dot{\tilde{w}}_{Rt})(z_{Hi} + \tilde{w}_{Ht}) f(z^m_{Rt}(z_{Hi})|z_{Hi}) + \int_{-\infty}^{z^m_{Rt}(z_{Hi})} \dot{\tilde{w}}_{Ht} f(z_{Ri}|z_{Hi}) dz_{Ri}\Big] dz_{Hi} \\ &= \tilde{\tilde{w}}_{Rt} + \tilde{\tilde{w}}_{Ht}.\end{aligned} \tag{A40}$$

Finally, it is easy to see that $\dot{z}^m_{Rt} = \dot{w}_{Ht} - \dot{w}_{Rt}$. Plugging in, we have

$$\begin{aligned}\dot{\bar{z}}_{Rt} &= (\dot{w}_{Rt} - \dot{w}_{Ht})\frac{M(z_{Ri} = z^m_{Rt}(z_{Hi}))}{M(z_{Ri} \geq z^m_{Rt}(z_{Hi}))}\left[\mathbb{E}\left(z^m_{Rt}(z_{Hi})\right) - \mathbb{E}\left(z_{Ri} \mid z_{Ri} \geq z^m_{Rt}(z_{Hi})\right)\right] \\ &+ (\dot{w}_{Rt} + \dot{w}_{Ht} - \dot{\mu}_{Ct})\frac{M(z_{Ci} = z^m_{Ct})}{M(z_{Ci} \geq z_{Ct})}\left[z^m_{Ct} - \mathbb{E}\left(z_{Ci} \mid z_{Ci} \geq z^m_{Ct}\right)\right],\end{aligned} \tag{A41}$$

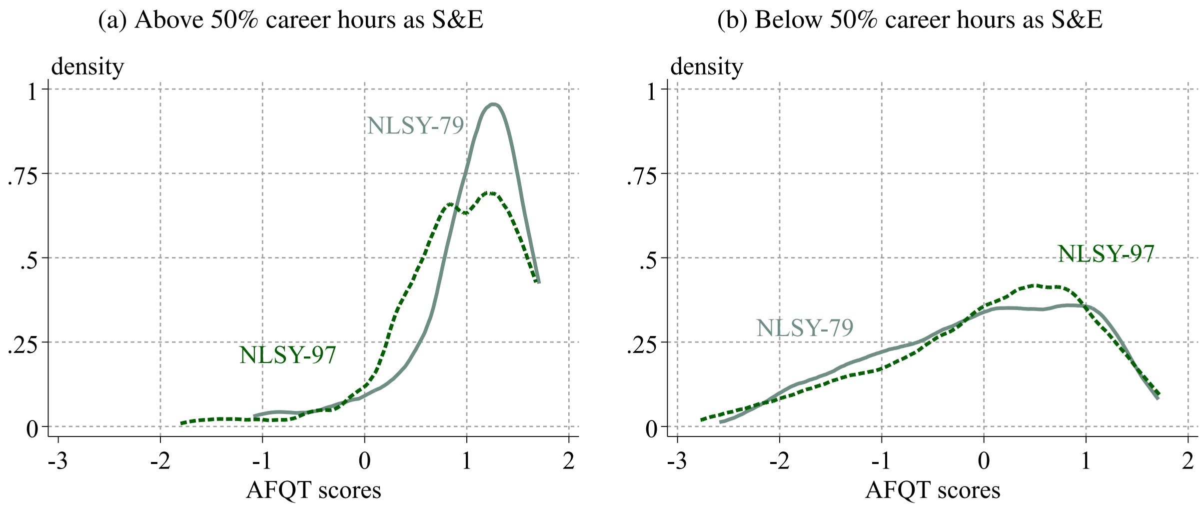
which is exactly equation (10).

B Data appendix

B.1 NLSY test scores: robustness to definition of researchers

Figure A1 reports the distribution of AFQT scores by NLSY cohorts for scientists and engineers (S&E) in Panel (a) vs. for other workers in Panel (b). The only difference from the results shown in Section I is the definition of S&E workers. In Section I, individuals are categorized as S&E if they have ever reported S&E as their occupations within the first 25 years of the survey; here, on the other hand, we alternatively define S&E as those with more than 50% of hour-weighted career within the first 25 years of the survey. As the figure shows, the results are robust with respect to these alternative definition. S&E workers' test scores shifts to the left and become more dispersed from the 79 to 97 cohorts, while we do not see such change for none-S&E workers.

Figure A1: Robustness to AFTQ score distribution by NLSY cohorts



Notes: Figure depicts AFQT score distributions in NLSY-79 (BLS, 2025) and NLSY-97 (BLS, 2026). Panel (a) shows those who have worked as researchers for more than 50% of their career working hours (within the first 25 years of the survey); Panel (b) shows those who have worked as researchers for less than 50% of their career working hours. Here, researchers are defined as those whose occupations are scientists, engineers, or related. The AFQT scores are taken from Altonji, Bharadwaj and Lange (2012) and standardized with respect to the mean and standard deviation of the NLSY-79 cohort. Density uses an Epanechnikov kernel with a 0.2 bandwidth.

B.2 Evidence for declining ability of researchers from patent data

If average researcher productivity is falling, one would expect this to show up in the cohort-specific distribution of productivity of innovators in publishing patents. To approach this, we use data on the universe of granted U.S. utility patents from the U.S. Patent and Trademark Office's (PTO) PatentsView database (USPTO, 2025). For each patent, we count the number of citations the patent receives over the 10-year window after the granting year. Windowing the number of citations ensures that results are not mechanically biased towards older patents. For each inventor, we restrict to patents granted in the first 10 years of the innovator's career, where this restriction is intended to minimize confounding effects due to inventor experience. The result is a measure of an innovator's patent productivity.

Table A1 regresses this measure on grant-year fixed effects and inventor cohort fixed effects, where we assign each inventor to a "birth" cohort based on the first year they are granted a patent. The cohort fixed effects thus capture the average patent productivity of inventors in that cohort relative to the 1980-1989 cohort, controlling for a time trend. The results show that later-cohorts have lower average patent productivity. The second column of Table A1 controls for technology-class (CPC)-specific time trends, and the third adds a quadratic in the average number of coauthors across an inventor's patents. The fourth through sixth columns repeat the same series of regressions, but using breakthrough patents as defined in Kelly et al. (2021) instead of log citations as the dependent variable. The lower productivity of later cohorts is consistent across all these specifications.

Table A1: Regression: cohort fixed effects (FEs) of innovator productivity

	log citations			breakthrough patents		
1990 - 1999 cohort FE	-0.140	-0.137	-0.134	-0.148	-0.047	-0.035
(standard error)	(0.003)	(0.003)	(0.003)	(0.022)	(0.022)	(0.022)
2000 - 2009 cohort FE	-0.362	-0.350	-0.343	-0.866	-0.643	-0.621
(standard error)	(0.004)	(0.004)	(0.004)	(0.027)	(0.026)	(0.026)
2010 - 2019 cohort FE	-0.567	-0.555	-0.543	-1.024	-0.747	-0.713
(standard error)	(0.004)	(0.004)	(0.004)	(0.028)	(0.027)	(0.027)
grant year FE	X			X		
grant year × CPC FE		X	X		X	X
avg. # of coauthors (quadratic)			X			X

Notes: Each column shows the result of regressing a measure of inventor productivity on cohort fixed effects, where the excluded category is the the 1980-1989 cohort. Cohorts are assigned based on the year an inventor first has a patent granted. Log citations are constructed by summing over the citations accumulated by an inventor from patents granted in a given year, counting only citations accumulated in the first 10 years after the patent is granted. The first column controls for year fixed effects; the second controls for year by technology class (CPC) fixed effects, where the technology class of a patent is defined as the first technology class listed on the patent; the third column adds a quadratic in the average number of coauthors on an inventor's patents in a given grant year. The fourth through sixth column repeat these regressions using the number of breakthrough patents as defined in Kelly et al. (2021) as a measure of inventor productivity. Source: U.S. PTO PatentsView, 1980-2019.

Though these results are suggestive of declining average ability among researchers, we want to emphasize that there are alternative interpretations. Because cohorts are assigned based on the year an inventor first shows up in our dataset, we cannot simultaneously control for both year trends and returns to experience; consequently, it could be that the lower productivity of later cohorts is due to lower experience rather than lower ability.[A2] Our truncation of an inventor's productivity at 10 years after their first grant partially addresses this; indeed Table A2 shows that average

[A2]This issue is of course not unique to our setting; in general it is impossible to separate cohort, time, and experience fixed effects when cohort and experience are both defined based on entry into the data set; see discussions in, e.g. Lagakos et al. (2018).

inventor experience within cohorts is not too different.[A3] If, however, the returns to experience are declining, e.g. because of the increasing burden of knowledge as argued in Jones (2009), our documented decline in cohort productivity could nevertheless partially reflect differences in cohort-specific experience.

Table A2: Mean inventor experience by cohort

cohort	1980-1989	1990-1999	2000-2009	2010-2019
mean inventor experience	2.214	2.314	2.382	1.935

Notes: Cohorts are assigned based on the year an inventor first has a patent granted, whereas inventor experience is grant year minus the year an inventor first has a granted patent.

B.3 Sample construction for NSCG and Decennial Census/ACS

When calculating aggregate trend of sectoral shares, $\{S_{Rt}, S_{R|Ct}\}$, we broadly use all observations in the NSCG (NCSES, 2021) and the ACS/Decennial Census (Ruggles et al., 2025) for individuals in the labor force who have valid information on education and occupations. The only additional restriction in the NSCG is that the primary work activity must be reported.

We apply stronger sample restrictions when calculating the moments on earning distributions, $\{E_{jt}, V_{jt}\}$ for each $j \in \{R, H\}$, and the NSCG panel moments $\{S_{R\to H}, S_{H\to R}, E_R^{R\to H}, E_H^{H\to R}\}$. The main purpose is to construct average hourly earnings that are comparable across individuals. Specifically, we restrict both the NSCG and the ACS/Decennial Census to individuals satisfying the following:

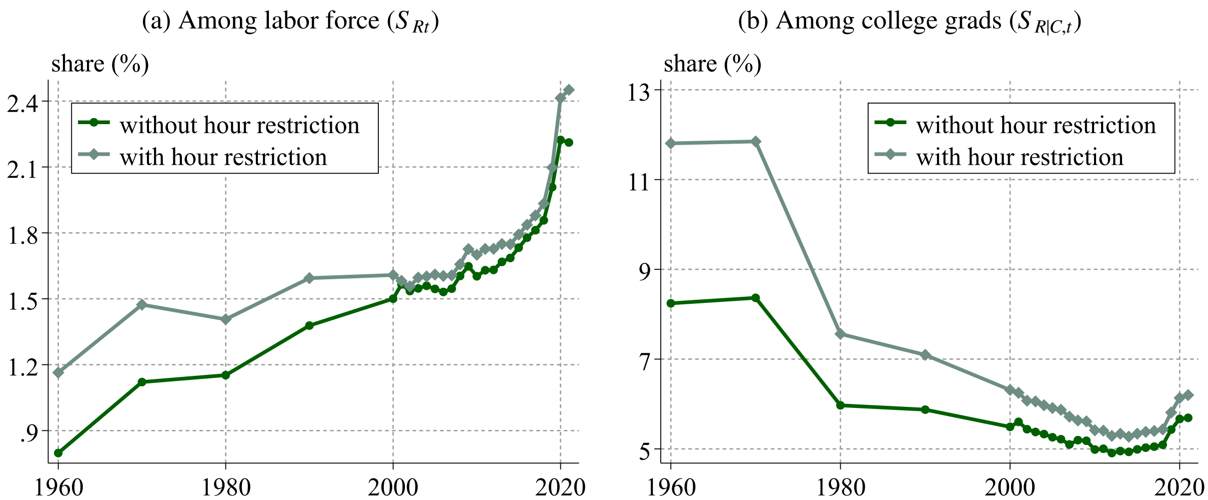
1. not missing any of salary, hours worked, weeks worked, gender, race, the hispanic dummy;
2. weekly hours worked less than 48,[A4]
3. annual weeks worked between 50 and 52,
4. wage greater than or equal to 0.95 times the federal minimum wage in the relevant year.

To demonstrate that this restricted sample reveals similar trends in researcher shares, Figure A2 compares the researcher shares, $\{S_{Rt}, S_{R|Ct}\}$, calculated using this restricted sample versus the broader one used in constructing the baseline share variables. The dark line depicts the shares in the broader sample, whereas the light line depicts those from the restricted sample. We demonstrate that the earning restrictions do not flip the overall trend of researcher shares. When calculating

[A3] With the caveat that the patent data starts in 1976 and so inventor experience is likely underestimated for earlier cohorts.

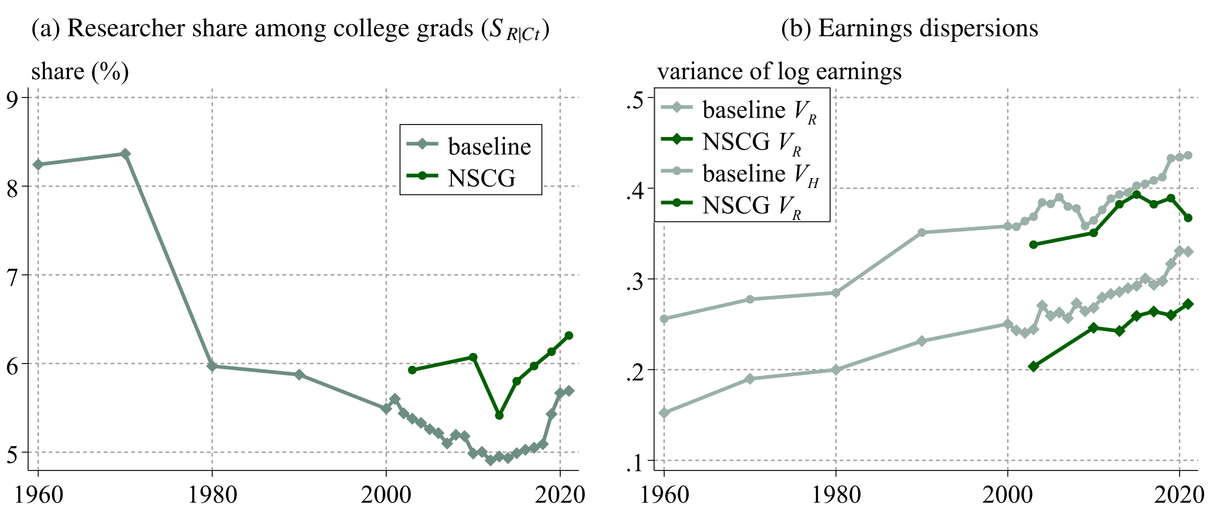
[A4] Weekly hours in some years are reported as intervals; for those years, we take the mid-point of the intervals.

Figure A2: Share of researchers across two samples



Notes: Numbers calculated by the authors using the NSCG and the Census/ACS.

Figure A3: Earnings dispersion: baseline vs. NSCG original



Notes: Numbers calculated by the authors using the NSCG and the Census/ACS.

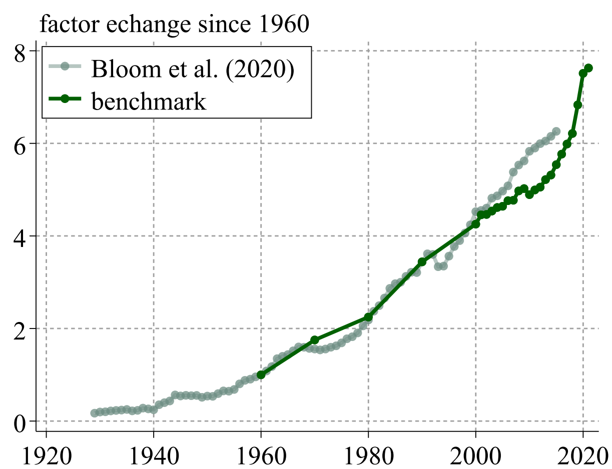
group-specific shares for foreign- vs. U.S.-born workers, we further restricted to the 1970-2021 sample where a consistent variable of workers' birthplace is observed.

We make use of the Census/ACS data for their larger sample size, longer sample period, and the information on non-college workers. For the years when NSCG is available, however, we can directly calculate, without imputing Census/ACS data, the variance of log earnings for the two types of college workers (V_{Rt} and V_{Ht}). Figure A3 demonstrates that our baseline imputed moments are similar to the directly-calculated moments from NSCG in our sample.

B.4 Comparing our researcher measure to Bloom et al. (2020)

We examine if our imputed trend is similar to other measures used in the literature. Figure A4 compares the factor change in the number of researchers, L_{Rt}, calculated in this paper with that in Bloom et al. (2020). The dark line depicts the authors' calculation using the NSCG and Census/ACS following equation (14). The light line, on the other hand, is Bloom et al. (2020)'s effective number of researchers.[A5] The figure shows that the growing share of researchers in our calculation is consistent with that in Bloom et al. (2020).

Figure A4: Number of researcher changes, authors' imputation vs. Bloom et al. (2020)



Notes: The dark line reflects the authors' calculation using the NSCG and the Census/ACS. The light line reflects Bloom et al. (2020)'s effective number of researchers. Both series are normalized to reflect factor changes from 1960.

[A5] Bloom et al. (2020) measure the effective number of researchers as the gross domestic investment in intellectual property products deflated by the nominal wage of males with a bachelor's or higher degree.

B.5 Changes in occupational researcher shares overtime

Recall that our baseline imputation procedure works as follows: first, for each occupation in the NSCG, we calculate the share of researchers using all NSCG survey years. We then match each Census/ACS occupation with an NSCG occupation and impute our share series into the Census/ACS.

This approximation relies on the assumption that within detailed occupations, the share of researchers are stable over time. To test whether our results are sensitive to this assumption, we recalculate the share of researchers using two different imputation procedures. The first calculates the average annual change rate of each NSCG occupation's researcher share and then imputes fitted values into the Decennial Census/ACS assuming constant within-occupation change rates.[A6] The second imputes to each Census/ACS year the occupational share of researchers from the *closest* NSCG year, e.g., 2003 for the 2001 ACS, 2010 for the 2011 ACS). Figure A5 shows the resulting time series. Though there are some level differences, the three imputation procedures result in very similar patterns.

This is because the occupational-level shares of researchers are indeed stable over time in our NSCG sample years. Figure A6 plots the occupational-level researcher shares, S_j^{NSCG}, in 2003 against their values in 2021. Each circle represents an occupation, and the circle size shows the number of college workers in the occupation, averaged over the two years. The circles are fairly close to the 45^o line. This suggests that the cross-time change of researcher shares within occupations is negligible compared with cross-sectional variation across occupations.

B.6 Are changes in earnings dispersion within or between occupations?

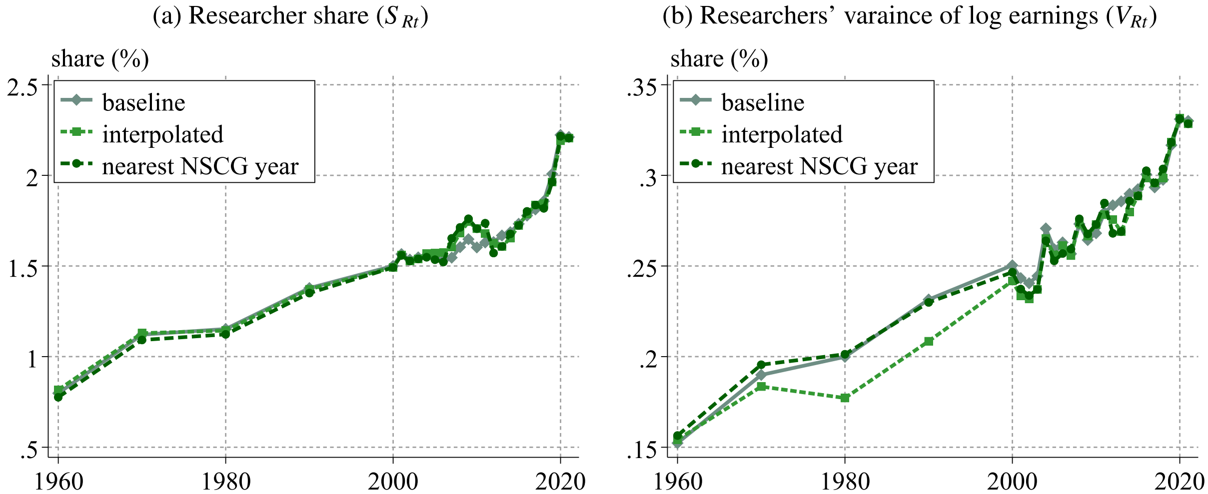
The stable occupational-level researcher shares also imply that the changes in earnings dispersion, ΔV_R and ΔV_H, over time are driven by the within-occupations instead of between occupations. To see this we decompose:

$$\Delta V_R = \underbrace{\Delta \sum_{j\in J} \left(\frac{L_{Rj}}{L_R}\right) V_{Rj}}_{\text{within-occupation}} + \underbrace{\Delta \sum_{j\in J} \left(\frac{L_{Rj}}{L_R}\right) \left(\tilde{E}_{Rj} - \tilde{E}_R\right)^2}_{\text{between-occupation}}, \tag{A42}$$

where $L_{Rj} = L_{Cj} S_j^{\text{NSCG}}$ is the occupation level number of researchers. The occupational level average and variance of log earnings for researchers, $\tilde{E}_{Rj}$ and V_{Rj} are imputed in the same manner

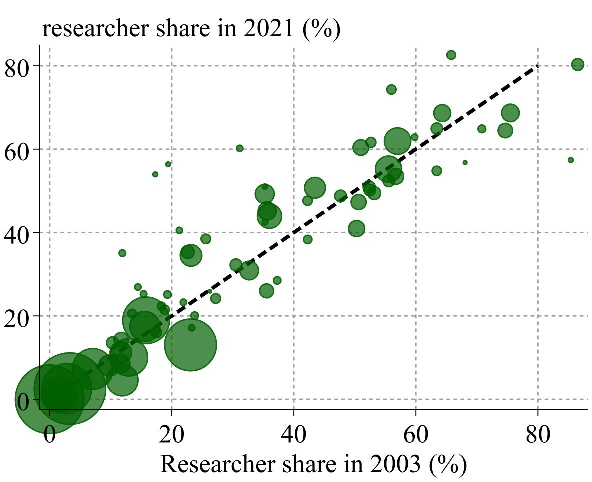
[A6] To be precise, we extrapolate occupational researcher shares in the Census/ACS before 2003, and interpolate in years between NSCG surveys, so that occupational researcher shares in NSCG years are exactly those in the NSCG. To keep shares bounded between 0 and 1, we construct researcher share change rates from the change rates of the number of researchers and the number of non-researchers within each occupation.

Figure A5: Comparing different imputation procedures



Notes: Numbers calculated by the authors using the NSCG and the Census/ACS..

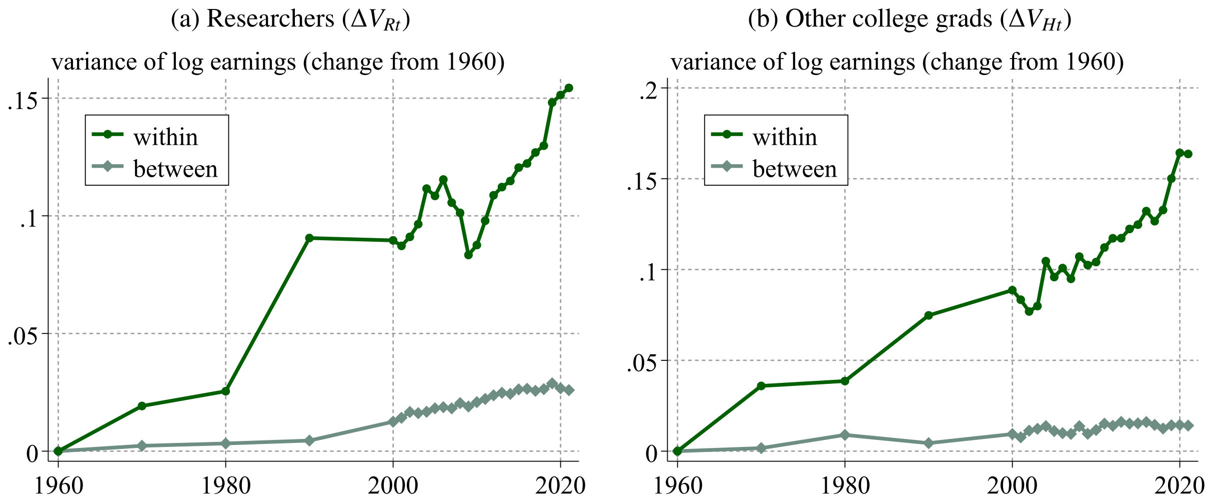
Figure A6: Occupational researcher share (S_j^{NSCG}) in 2003 against 2021



Notes: Each circle refers to an occupation in NSCG; the circle size is proportional to the number of workers in the occupation. Source: NSCG.

as in the main text, where the imputation is applied within each occupation. We also apply the same decomposition to ΔV_H. The decomposition are reported in Figure A7, for ΔV_R in Panel (a) and ΔV_H in Panel (b). We find that almost all changes in the variances comes from the within-occupation component.

Figure A7: Changing variance of log earnings: within vs. between occupations



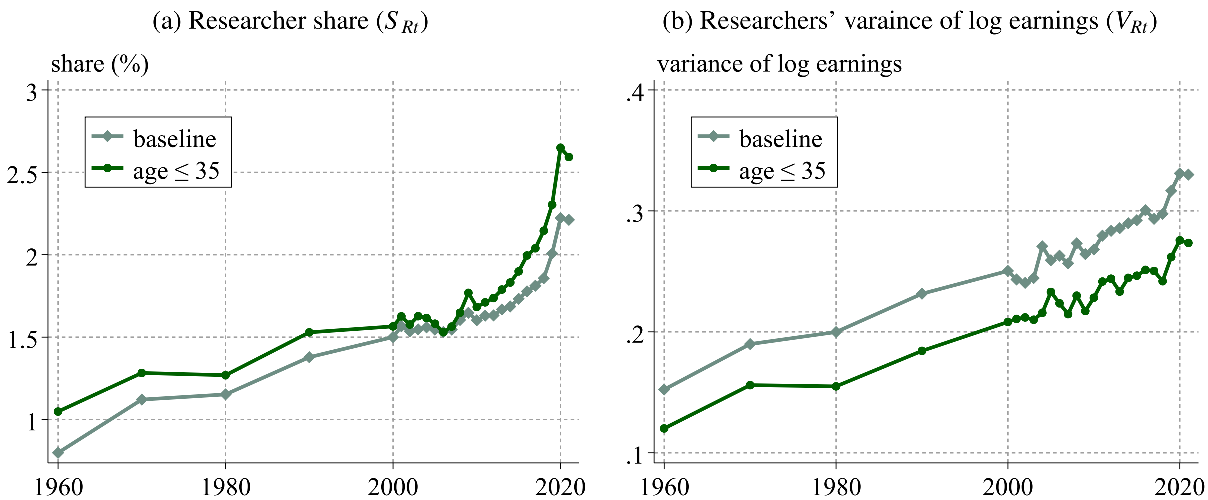
Notes: Numbers calculated by the authors using the NSCG and the Census/ACS.

B.7 Earnings moments among young workers

Even within an occupation, the earnings of researchers could be increasingly dispersed without the composition effect: older researchers are getting rents from past innovations, while younger researchers have a hard time coming up with breakthrough ideas. Therefore, the increasing dispersion could be reflecting the aging of researchers.

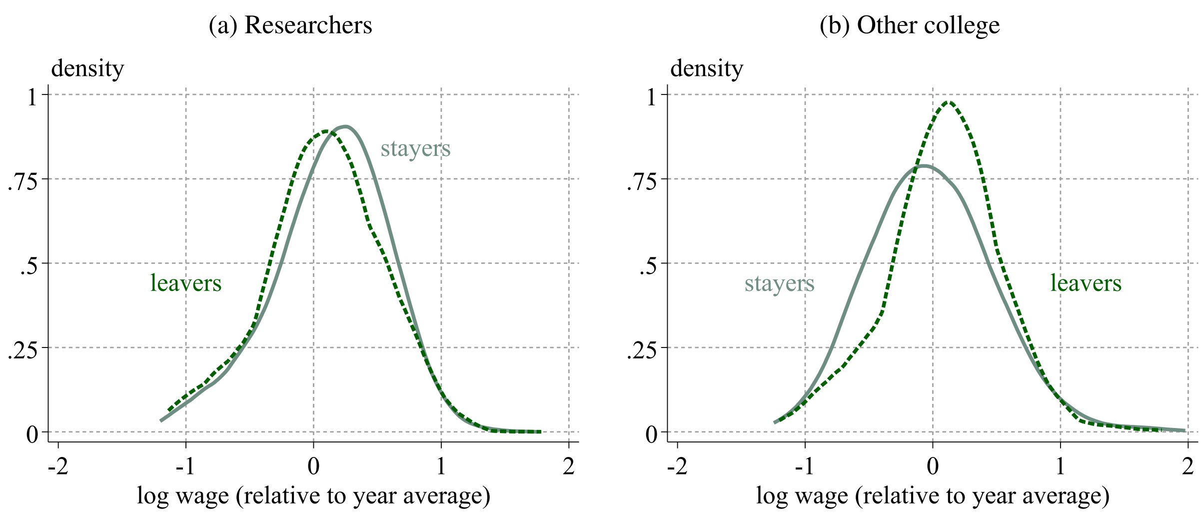
As a robustness check, we show that, when restricting our sample to young researchers, the share and earnings dispersion of researchers increase at a similar magnitude as in our baseline. Figure A8 shows the comparison, while we restrict the sample to researchers with an age below 35. Panel (a) plots the sequence of the researcher's share in the labor force; Panel (b) plots the variance of log earnings among researchers. We find that the baseline trends hold robustly among young researchers. This supports our conjecture that the trends are driven by self-selection of abilities across cohorts, instead of the aging of researchers. Figure A9 shows that the panel moments are also robust among young workers.

Figure A8: Researcher share and earnings dispersion: baseline vs. age ≤ 35



Notes: Numbers calculated by the authors using the NSCG and the Census/ACS.

Figure A9: Earning distributions in the NSCG 2010-2015 panel data, age ≤ 35



Notes: Numbers calculated by the authors using the NSCG.

B.8 Longitudinal moments in the NSCG and CPS

In Section III.D, we show that the share of movers among researchers is 20.3%, much higher than the 1.2% for non-researchers. To examine the external validity of this finding, we calculate a similar moment in the Current Population Survey (CPS).[A7] Recall that our definition of a researcher in the NSCG is an individual who reports working in a science and engineering (S&E) occupation (or related) and whose primary working activity is research or development. Though the CPS has an occupation code whose coding is similar to the NSCG, it lacks information about primary work activity, and so we cannot precisely replicate our researcher definition in the CPS. To make progress, we instead focus on a superset of our researchers, namely workers in S&E occupations.

Table A3: Transition rates between S&E and non-S&E occupations, NSCG vs. CPS

NSCG

	from	
to	S&E	not S&E
S&E	0.841	0.073
not S&E	0.159	0.927

CPS

	from	
to	S&E	not S&E
S&E	0.726	0.052
not S&E	0.274	0.948

Table A3 implements this comparison by showing transition rates between S&E and non S&E occupations in the NSCG vs. the CPS.[A8] The share of movers among researchers (equal to the bottom left element of the transition matrices) is similarly high in both the NSCG and the CPS, whereas the share of movers among non-researcher college graduates (equal to the top right element of the transition matrices) is similarly low in both the NSCG and the CPS. Note that the CPS transition matrix is annual, whereas the NSCG transition matrix combines moves from 2010 to 2013 and moves from 2013 to 2015.[A9] This caveat aside, this exercise supports the external validity of our finding that moving is much more common among researchers than among non-researcher college graduates.

The other finding from our longitudinal analysis of the NSCG is that researchers moving to non-research tend to have lower earnings than researchers who do not move, whereas the opposite pattern is observed for non-researchers moving to research. Table A4 calculates these moments in the CPS, using the same sectoral definitions as in the previous paragraph. The results are qualitatively consistent with our findings from the NSCG: workers moving from S&E to non-S&E

[A7] Data extracted and downloaded from IPUMS CPS (Flood et al., 2025).

[A8] Note that the NSCG portion of the table does not exactly replicate what we did in the main text because we do not apply the harmonizations needed to map to the Census/ACS in this subsection.

[A9] We could make a more direct comparison by e.g. inferring the biennial transition matrix in the CPS assuming that it is generated by a first-order Markov process, but this assumption is likely a poor approximation to the data—indeed our model does not satisfy it.

occupations have lower earnings than those who do not move, whereas workers moving from non-S&E to S&E occupations have higher earnings than those who do not move.

Table A4: Differences in mean log earnings between movers and stayers, CPS

	mean log earning	
	S&E	not S&E
stayer	7.659	7.223
mover	7.494	7.420

B.9 Researcher shares beyond the United States

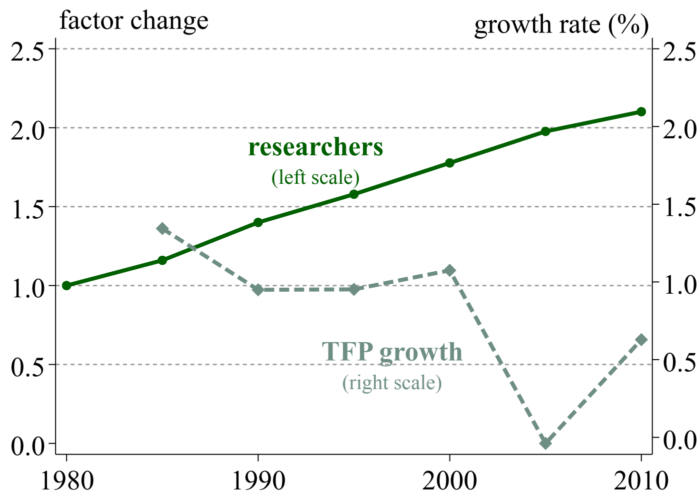
We want to highlight that the facts documented in Section III.C are valid in other developed countries beyond the United States, suggesting that the forces we highlight are not unique to the U.S. context. In this section, we present the cross-time and cross-country patterns of R&D workforces for 18 countries whose time series of annual R&D personnel from 1980 to 2010 are available in OECD Statistics (OECD, 2026*b*).[A10] We calculate the total number of researchers, the population share of researchers, and the share of researchers among college graduates for each country and each five-year interval from 1980 to 2010. The number of researchers is taken from OECD Statistics and is defined as the full-time equivalent number of persons engaged in R&D.[A11] Data for college graduates are taken from Barro and Lee (2013), and is defined as the population with a complete tertiary education. Data for GDP per person are taken from the Penn World Table (PWT) 9.1 (Feenstra, Inklaar and Timmer, 2015). Data for TFP growth rates are also from OECD Statistics (OECD, 2026*a*).

We first show that average researcher productivity has also been declining over time for the 18 OECD countries. To see this, we take the population-weighted average over the 18 countries and examine the time trends of average annual TFP growth and the factor change in the number of researchers since 1980. The results are presented in Figure A10. While the number of researchers doubled from 1980 to 2010, the average annual growth rate of TFP was flat at around 1% before 2000 and dropped during the 2003-2007 interval. This implies that the average productivity of researchers has fallen to less than half of its 1980 level.

[A10]The 18 countries are: Australia, Austria, Belgium, Canada, Denmark, France, Germany, Hungary, Iceland, Ireland, Italy, Japan, Netherlands, Norway, Spain, Sweden, United Kingdom, and the United States. Data was downloaded from the *OECD Statistics* website: https://stats.oecd.org/. The website has now been replaced by *OECD Data Explorer* (OECD, 2026*a*,*b*).

[A11]The methodology for measuring R&D personnel in this data set is documented in the OECD's Frascati Manual: http://www.oecd.org/sti/inno/Frascati-Manual.htm.

Figure A10: Factor changes in researchers vs. TFP growth, average over 18 OECD countries



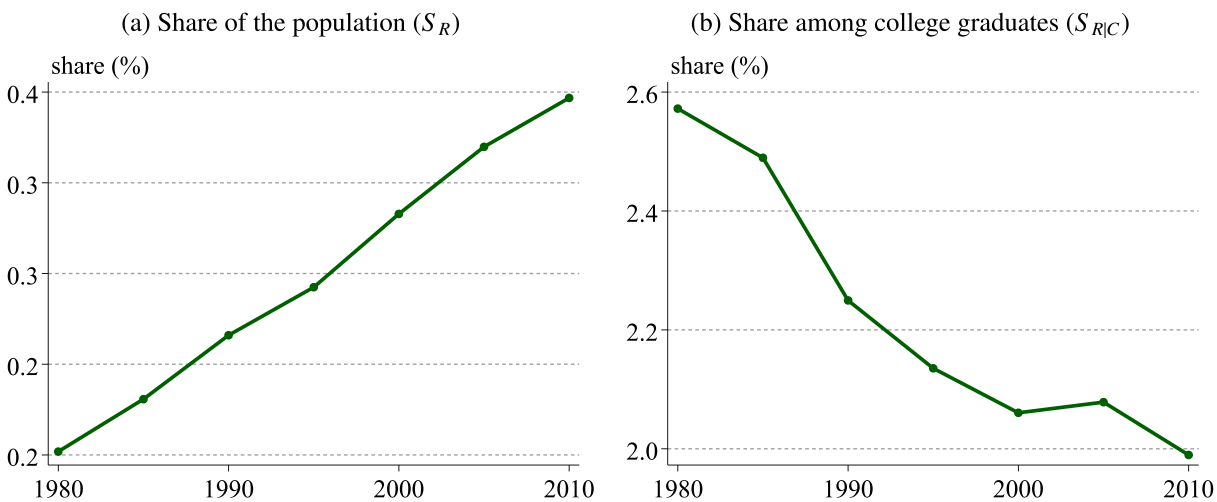
Notes: The solid line refers to the left vertical axis, depicting the factor change of researchers since 1980. The dashed line refers to the right vertical axis, depicting the average annual growth rate of TFP. Data are reported in five-year intervals. We match the scales of the two vertical axes so that proportional changes of the two series are comparable. Data source: see Section B.9.

We next examine the trend of average researcher-to-population shares for the 18 OECD countries. Panel (a) of Figure A11 shows the population-weighted average share of researchers (S_{Rt}) over time, and Panel (b) shows the researcher shares relative to the college-educated population ($S_{R|C,t}$).[A12] The direction of both trends is consistent with their U.S. counterparts as presented in the previous section. The overall population share of researchers doubled from 1980 to 2010, and the share among college graduates in 2010 is around four-fifths of its 1980 level.

Figure A12 shows the cross-country analog for the evolution of researcher shares along the path of development. Panel (a) plots the overall share of researchers (S_{Rt}) by log GDP per person for each country-year observation. It shows a clear positive correlation between researcher share and per-person income. We further project the researcher shares on log income per person, controlling for year-fixed effects (regressions are weighted by population size). Panel (b) plots the researcher share among college graduates ($S_{R|C,t}$) by per capita log GDP, showing an analogous pattern to panel (b) of Figure A11, projecting the shares on log GDP per person controlling for year fixed effects. We conclude that the cross-time and cross-country patterns are both consistent with those observed in the United States.

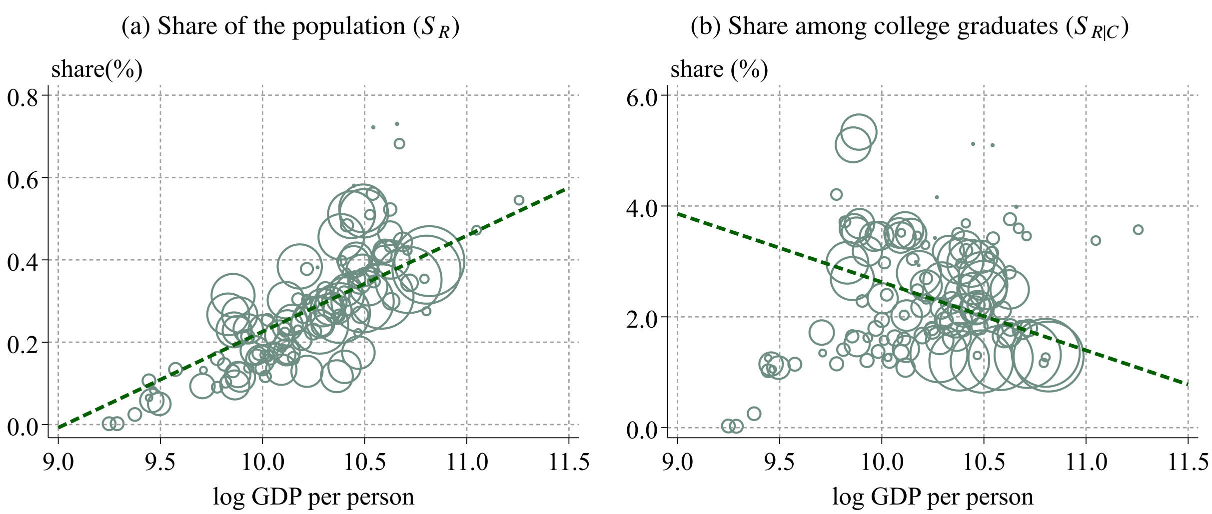
[A12]Note that here we calculate the population share, instead of the labor force share of researchers. This is because Barro and Lee (2013) only provides the population shares for education groups. To further refine our measurement, one can replace it with data from the International Labor Organisation (ILO) to obtain employment data by education level. This way, we will be able to measure the employment share of researchers.

Figure A11: Time trend of researcher shares, average over 18 OECD countries



Notes: Panel (a) depicts the share of researchers relative to the total population (S_R). Panel (b) depicts the share of researchers relative to college graduates ($S_{R|C}$). Data are reported in five-year intervals. Data source: see Section B.9.

Figure A12: Researcher shares vs. income per person, 18 OECD countries, 1980-2010



Notes: Each circle depicts the share of researchers by its log income per person for a country-year observation (18 countries, each has 7 observations from 1980 to 2010 in five-year frequency). The sizes of the circles represent countries' relative populations. The dashed lines depict the fitted value of population-weighted OLS regression. Panel (a) reports the overall share of researchers, while panel (b) reports the share of researchers among college graduates. Data source: see Section B.9.

C Estimation and computation appendix

C.1 Summary of computational algorithm

Recall that the objective function is:

$$\hat{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}} \left[\boldsymbol{M}_2(\boldsymbol{p}) - \boldsymbol{D}_2\right]' \boldsymbol{W} \left[\boldsymbol{M}_2(\boldsymbol{p}) - \boldsymbol{D}_2\right], \quad \text{subject to} \;\; \boldsymbol{M}_1(\boldsymbol{p}) = \boldsymbol{D}_1. \tag{A43}$$

Let $\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2)$, with $\boldsymbol{p}_1 = \boldsymbol{M}_1^{-1}(\boldsymbol{D}_1\,;\,\boldsymbol{p}_2)$. The function $\boldsymbol{M}_1^{-1}$ is known given a $\boldsymbol{p}_2$, and $\boldsymbol{D}_1$ comes directly from the data; the algorithm thus just needs to solve for $\boldsymbol{p}_2$. To do so, we start with an initial guess $\boldsymbol{p}_{2,0}$, a convergence criterion, an updating rule, and two $N \times S$ matrices of simulates $\boldsymbol{X}_R, \boldsymbol{X}_H$ drawn iid from a standard normal distribution. The algorithm proceeds as follows:

1. Use $\boldsymbol{p}_{2,0}$ to construct $N \times S$ matrices of simulated sector-specific abilities $\boldsymbol{Z}_R, \boldsymbol{Z}_H$ by applying the Gaussian copula $C(\boldsymbol{X}_R, \boldsymbol{X}_H)$, where for an element (x_R, x_H) the copula returns $(\sigma_R x_R, \sigma_H(\rho x_R + \sqrt{1-\rho^2}x_H))$ and C just applies this function element by element. As can easily be confirmed, the result of this procedure are draws from a multivariate normal distribution with marginal means 0 and variance-covariance matrix defined using σ_R, σ_H, ρ.
2. Given $\boldsymbol{p}_1 = \boldsymbol{M}_1^{-1}(\boldsymbol{D}_1\,;\,\boldsymbol{p}_2)$, solve the model S times by assigning the N simulated sector-specific abilities from a column of $\boldsymbol{Z}_R, \boldsymbol{Z}_H$ to sectors R and H, computing the sector-specific change of the variance of log earnings and the sector-specific difference in the mean log earnings of movers and stayers. Average the S resulting set of moments.
3. Add the change in the variance of z_C among college graduates to the sector-specific change of variance from the previous step; given $\boldsymbol{p}$ this is just the variance of a truncated normal distribution with known truncation point and can therefore be solved without simulation.
4. Terminate if the convergence criterion is met; otherwise apply the updating rule and restart from 1.

In practice, we search over the Cholesky factors of the variance-covariance matrix implied by a given $\boldsymbol{p}_2$. This improves the estimation algorithm considerably, first because we do not need to check if the variance-covariance matrix is positive-definite at each iteration and second because it avoids problems associated with approaching the $\rho = 1$ boundary. We generally use $N = 2143, S = 14$ for estimation.[A13] The termination criteria are a) the decrease in the objective function between

[A13]These are picked using the elbow rule on the variance of simulated moments across many seeds i.e. different $\boldsymbol{X}_R, \boldsymbol{X}_H$. We use higher N for sensitivity calculations and to generate figures, and lower N for the bootstrapping procedure.

iterations is less than or equal to 1×10^{-20}, b) the relative change in the Euclidean norm of the parameters is less than 1×10^{-4}, c) the algorithm has reached 7500 iterations. The estimation is done in two parts: first, we apply the Dividing Rectangles algorithm (a global solver) from Julia's `OptimizationNLopt` package; once the algorithm has terminated, the solution is used as the initial guess for the Nelder Mead algorithm (a local solver) from `OptimizationNLopt`. The algorithm usually converges in under 8000 iterations, with the bulk of those consumed by the global solver.

C.2 Constructing numerically-targeted moments in simulated model

We construct model counterparts for numerically-targeted moments, $\boldsymbol{M}_2(\boldsymbol{p})$, through a combination of analytical solutions and simulation. Because the random variable $z_C \mid s = 1$ is a truncated normal, its first two moments are known functions of parameters. Moreover, the distribution of $z_C \mid s = 1$ does not depend on whether an individual chooses to become a researcher or not. As a result, we do not need to draw z_C in our simulations. Specifically, given a simulation of z_R and z_H, we can calculate mean log earnings by

$$E_j = w_j + \underbrace{\mathbb{E}\left(z_{Ci} \mid s = 1\right)}_{\text{analytical}} + \underbrace{\mathbb{E}\Big(z_{ji} \mid s = 1, j = j\Big)}_{\text{simulated}}, \tag{A44}$$

for each $j \in \{R, H\}$. Likewise, we construct model counterparts for ΔV_R and ΔV_H by

$$\Delta V_j = \underbrace{\Delta \mathrm{Var}\left(z_{Ci} \mid s = 1\right)}_{\text{analytical}} + \underbrace{\Delta \mathrm{Var}\Big(z_{ji} \mid s = 1, j = j\Big)}_{\text{simulated}}, \tag{A45}$$

for each $j \in \{R, H\}$. To construct model counterparts for longitudinal moments, $E_R^{R\to H}$ and $E_H^{H\to R}$, we first determine a set of sectoral movers in our model. To do so, we assume that movers are those on the margin of choosing the other sector, i.e., those with weaker comparative advantage in their sector. To be precise, we pick the thresholds, $\overline{\alpha} > \underline{\alpha}$, such that the sectoral shares of movers in the 2010-2015 NSCG panel are exactly replicated:

$$\begin{aligned} S_{R\to H} &= \Pr\left(\tilde{w}_H - \tilde{w}_R < z_{Ri} - z_{Hi} < \overline{\alpha}\right) \\ S_{H\to R} &= \Pr\Big(\underline{\alpha} < z_{Ri} - z_{Hi} < \tilde{w}_H - \tilde{w}_R\Big), \end{aligned} \tag{A46}$$

as listed in Panel (a) of Table 2. Given the thresholds, the mean log earnings of movers relative to stayers in the model are then calculated by simulation:

$$\begin{aligned} E_R^{R\to H} &= \mathbb{E}\left(e_{Ri} \mid \tilde{w}_H - \tilde{w}_R < z_{Ri} - z_{Hi} < \overline{\alpha}\right) - \mathbb{E}\left(e_{Ri} \mid z_{Ri} - z_{Hi} > \overline{\alpha}\right) \\ E_H^{H\to R} &= \mathbb{E}\Big(e_{Hi} \mid \underline{\alpha} < z_{Ri} - z_{Hi} < \tilde{w}_H - \tilde{w}_R\Big) - \mathbb{E}\Big(e_{Hi} \mid z_{Ri} - z_{Hi} < \underline{\alpha}\Big), \end{aligned} \tag{A47}$$

We do this exercise using the 2021 relative sectoral returns ($\tilde{w}_H - \tilde{w}_R$) because the panel years are much closer to the end of our sample period than to the initial year.

Summing up, our estimation algorithm consists of the following steps. First, take a candidate minimizer $\boldsymbol{p}$ such that the moments in Panel (a) of Table 2 are exactly matched, i.e., $\boldsymbol{M}_1(\boldsymbol{p}) = \boldsymbol{D}_1$. Second, simulate the model and calculate the model counterpart of moments listed in Panel (b) of Table 2, i.e., $\boldsymbol{M}_2(\boldsymbol{p})$ and calculate its distance with $\boldsymbol{D}_2$. Third, search for the estimate $\hat{\boldsymbol{p}}$, that minimizes the distance between model and data moments as defined in equation (23).[A14]

C.3 Constructing the bootstrapped confidence interval

To construct the bootstrapped confidence interval for the estimated parameters and the ratio change in average ability, we first create 1000 bootstrap samples by resampling with replacement from the cleaned sample that is used for the baseline analysis. The resampling is stratified by year and by the subsample that is used for earning moments, i.e., those observations with work hour information and that satisfy our sample selection for earnings. We then re-estimate the model targeting each of the 1000 bootstrapped moments. This gives us 1000 estimates for each parameter. We then take their 2.5 and 97.5 percentiles to form the 95% confidence intervals. Panel (a) of Figure A13 depicts the distribution of the 1000 bootstrapped estimates for $\Delta \ln \overline{Z}_R$. The black dashed line indicates the point estimate of 52.6%, whereas the green solid line indicates the lower and upper bounds of the confidence interval, respectively. Panel (b) plots the upper and lower bounds of the confidence interval against the total number of bootstrapped samples used. This ensures that our number of bootstraps is large enough so that the interval converges.

C.4 Trend in earnings dispersion, data vs. model

To provide additional validation and illustration, we also plot the model-implied sequence of earnings dispersion in the two high-skill sectors, $\{V_{Rt}, V_{Ht}\}$ in Figure A14, which is untargeted save for the first and last points.

C.5 Sensitivity of estimated parameters with respect to moments

To further investigate the mapping from data moments to model parameters imposed by our model, we implement the sensitivity measure defined in Andrews, Gentzkow and Shapiro (2017). Sensi-

[A14]In practice, the algorithm searches over a transformation of $\boldsymbol{p}$, namely the elements of the Cholesky factor of the variance-covariance matrix implied by $\boldsymbol{p}$. This elides the need to check positive-definiteness for each $\boldsymbol{p}$, improving speed and stability.

Figure A13: Bootstraped distribution for $\overline{Z}_{R,2021}/\overline{Z}_{R,1960} - 1$ and its 95% confidence interval (CI)

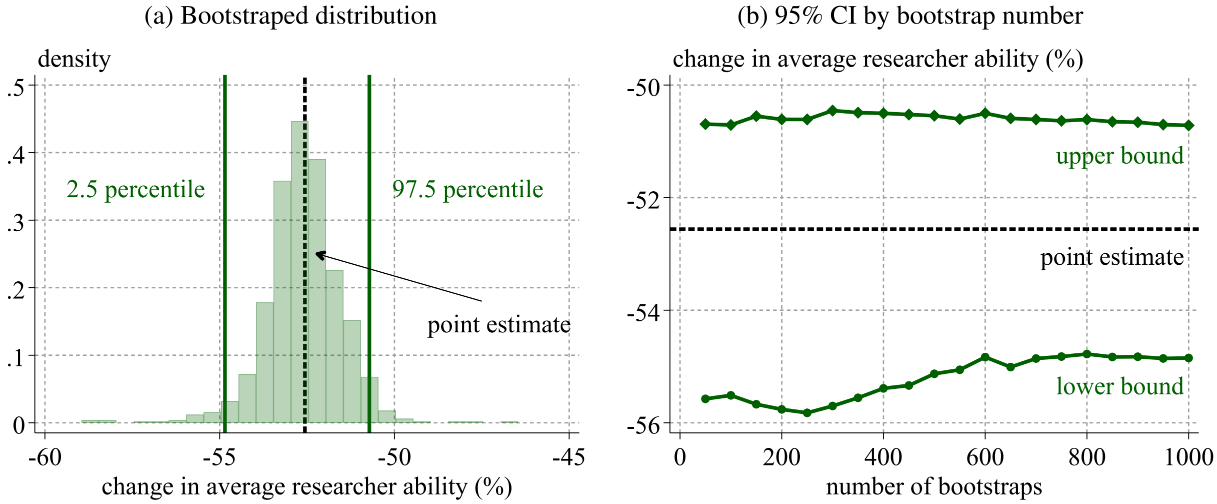
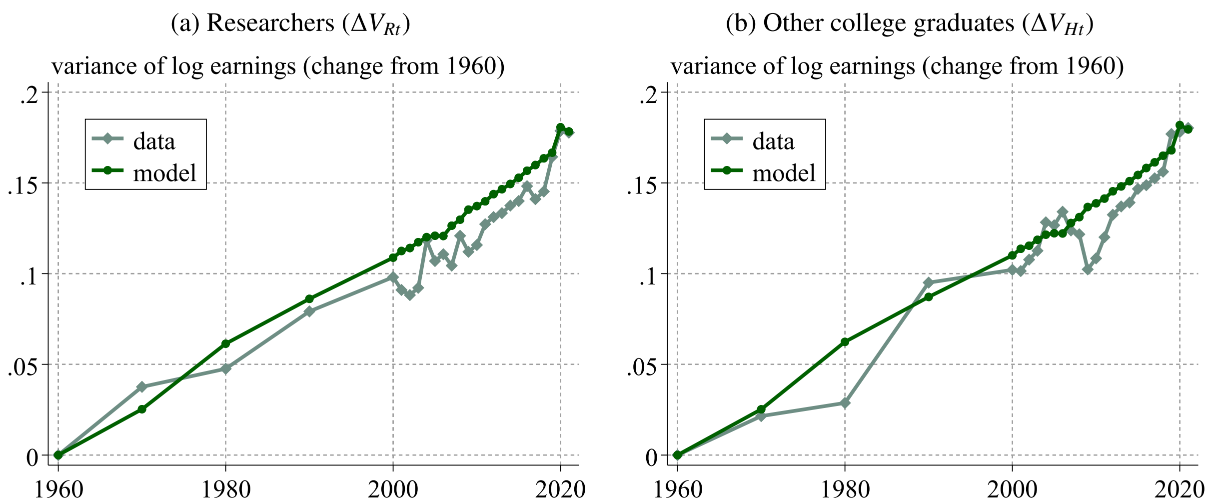


Figure A14: Change in earnings dispersion, model vs. data



Notes: Panel (a) plots the change in variance of log earnings among researchers from 1960. Panel (b) plots that of other college graduates. In each panel, the dark line depicts the model-predicted value, the light line depicts the data value calculated using NSCG and ACS/Census.

tivity Λ is defined as:

$$\Lambda = \left[G(\hat{\boldsymbol{p}})'\boldsymbol{W}G(\hat{\boldsymbol{p}})\right]^{-1} G(\hat{\boldsymbol{p}})'\boldsymbol{W}, \tag{A48}$$

with $G(\hat{\boldsymbol{p}})$ the Jacobian of the vector of targeted model moments evaluated at the parameter estimate $\hat{\boldsymbol{p}}$ and $\boldsymbol{W}$ the weight matrix used in the estimation. It is termed sensitivity because it captures the negative of first-order asymptotic bias in parameter estimates with respect to a local perturbation in the assumptions on the data-generating process. Intuitively, an element (i, j) of Λ corresponds to the sensitivity of parameter i with respect to a change in moment j.

Table A5 shows the elements of Λ under our estimated parameter vector. Each row reports the sensitivity of a given parameter to the numerically approached targeted moments. The elements of Λ are scaled so that the numbers are interpretable as the asymptotic negative bias in a parameter given a 1% change in a particular moment, shown in units of percent change from the point estimates.[A15] In line with our discussion in Section IV.C, the standard deviation of sector-neutral ability (σ_C) is particularly sensitive to the change in the variance of log earnings among researchers and among non-researcher college graduates. The sensitivity to the panel moments is highest for the distributional parameters of the sector-specific ability (σ_R, σ_H, ρ); this is again in line with our discussion in IV.C.

We can further calculate the sensitivity of the estimated change in ability among researchers, which by the delta method is just $C(\hat{\boldsymbol{p}}) \times \Lambda$ with $C(\hat{\boldsymbol{p}})$ the Jacobian of the change in ability among researchers with respect to the parameter vector, evaluated at the estimates. The bottom row of Table A5 shows the resulting vector.[A16] The estimated change in ability among researchers is particularly sensitive to the sector-specific changes in the variance of log earnings; this follows because the model infers almost the entire decline in ability as coming from the education margin of the model, the strength of which is controlled by σ_C, which as discussed earlier is most sensitive by said moments.

We want to emphasize that the sensitivity is evaluated locally: if the data moments were such that the model attributed a greater role to the sectoral choice channel, then the sensitivities on the panel moments and the change in the variance of log earnings among researchers would be magnified. The results should thus not be interpreted as saying that moments other than the change in the variance of log earnings among non-researcher college graduates are irrelevant in pinning down the change in ability among researchers; indeed, these sensitivities are non-negligible even locally.

D Estimation details for main text Section VI.A

[A15] Because the objective function used to estimate the model belongs to the class of classical minimum distance estimators as defined in Andrews, Gentzkow and Shapiro (2017), the most relevant set of perturbations to consider are additive shifts to the moment functions, in which case the asymptotic negative bias under a given set of alternative assumptions resulting in additive shifts to the moment functions η is simply $\eta \times \Lambda$. Note that we end up with the asymptotic negative bias because our objective function is defined as model minus data instead of data minus model as in Andrews, Gentzkow and Shapiro (2017).

[A16] Note that these are again shown as percent change from the point estimate; accordingly a larger positive number implies a stronger decrease in average ability among researchers, and vice versa.

Table A5: Sensitivity of parameters to numerically-approached moments (Λ)

	ΔV_R	ΔV_H	$E_R^{R\to H}$	$E_H^{H\to R}$
σ_C	1.411	1.353	0.076	0.037
σ_R	−0.581	−0.559	0.579	−0.223
σ_H	0.669	0.645	−1.566	0.563
ρ	−0.105	−0.100	−0.245	0.079
$\Delta \ln \overline{Z}_R$	0.790	0.758	0.060	0.015

Notes: Each cell reports an element of the sensitivity matrix Λ as defined in equation (A48). The value measures the sensitivity of an estimated model parameter (row) to a moment targeted in the estimation (column). The last row shows the sensitivity of the estimated change of ability among researchers between 1960 and 2021 ($\Delta\overline{Z}_R$) to each targeted moment. All elements are divided by 100 so that they capture the sensitivity of a parameter with respect to a 1% change in a given moment.

D.1 Robustness to residualizing earnings

In this subsection, we residualize the log hourly earnings in both the NSCG and the ACS/Census with respect to demographic controls: the full interaction of gender, race, and Hispanic identifier. Because we observe education in ACS, we interact the demographic controls with the college indicator and the survey year; likewise, because we interact the controls with the researcher indicator and the survey year in the NSCG. We then repeat our baseline estimation using the average residualized log hourly earnings.

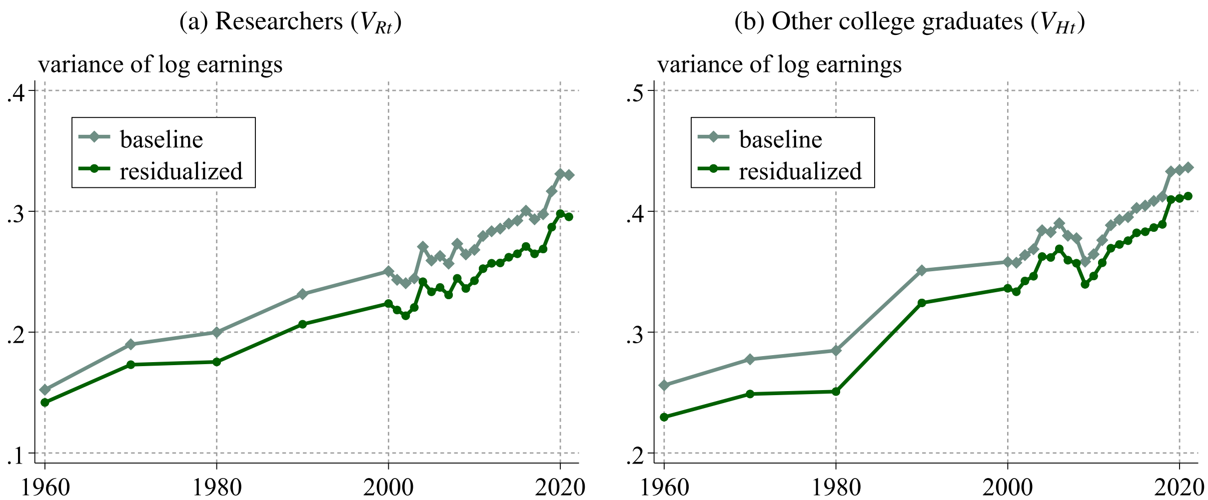
We first show that the earning-related targeted moments are virtually unchanged with vs. without residualization. Figure A15 compares the trends in variance of log earnings for researchers in Panel (a) and for other college graduates in Panel (b). While residualization slightly lowers the dispersions for each year, the changes are virtually the same over our sample period. Figure A16 shows that leavers' vs. stayers' residualized earnings distribution in the 2010-2015 NSCG panel. It shows the same pattern as Figure 7 in Section III.D: leavers earn 5.2% less on average in the research sector, but they earn 13.4% more among other college graduates. Using the residual earning moments as the target, the indirect inference indicates a 52.5% decline in the average ability of researchers from 1960 to 2021. This result is virtually the same as our baseline.

D.2 Re-estimating model allowing for group-specific wedges

D.2.1 Estimation results

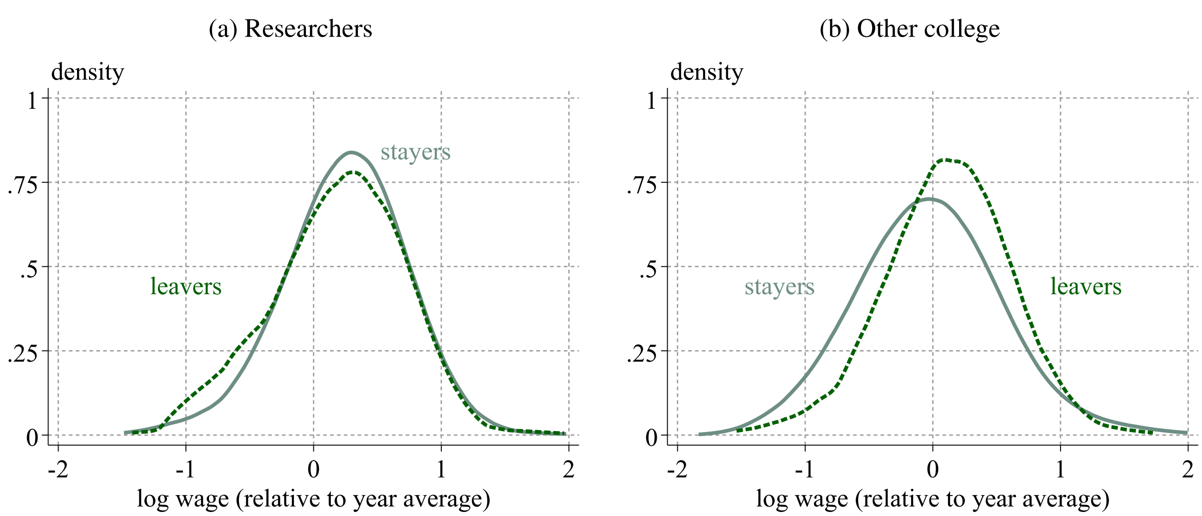
To estimate the extended model for Section VI.A and Section VI.A, we target data moments in the baseline plus the group-specific shares, $\{S_{Cg}, S_{R|Cg}, \Delta S_{Cg}, \Delta S_{R|Cg}\}$, while taking the unconditional group composition $\{S_g, \Delta S_g\}$ as exogenously given from the data. Table A6 reports these moments

Figure A15: Variance of log earnings: baseline vs. residualized



Notes: Numbers calculated by the authors using the NSCG and the Census/ACS.

Figure A16: Earning distributions in NSCG panel: residualized



Notes: The Epanechnikov kernel with 0.2 bandwidth is used for the density. Data source: NSCG 2010-2015 panel.

in the data. We treat them as exactly-matched moments, $\boldsymbol{D}_1 = \boldsymbol{M}_1(\boldsymbol{p})$, in the estimation algorithm. Table A7 reports the resulting estimated labor wedges and their changes over the sample period $\{\tau_{Cg}, \tau_{Rg} - \tau_{Hg}, \Delta\tau_{Cg}, \Delta(\tau_{Rg} - \tau_{Hg})\}$. Note that we can only identify the total number of groups minus 1 sector-specific wedges; accordingly, we normalize the wedges of white-male workers in the model with race and gender groups and those of U.S.-born workers in the model with immigrants.

D.2.2 Allowing for group-specific differences in ability

Our baseline results allowing for sector-specific demographic compositions restrict the mean of ability to be the same across groups. In this subsection, we consider the opposite extreme case: assuming no group-specific distortions in the returns to ability and thereby attributing all cross-group gaps in researcher shares to ability differences.

Given the estimated parameters from above, we can further incorporate group-by-sector-specific average earnings and their changes, $\{E_{Rg}, E_{Hg}, \Delta E_{Rg}, \Delta E_{Hg}\}$, to decompose the estimated wedges into earnings and non-earnings components:

$$\tau_{jg} = \tau^w_{jg} + \tau^\mu_{jg}, \tag{A49}$$

so that earnings in the model become:

$$e_{ji} = w_j + z_{Ci} + z_{ji} + \tau^w_{jg}. \tag{A50}$$

For the analysis in Sections VI.A and VI.A, estimating the change in ability does not require us to separate these components. When considering the case where we attribute all cross-group differences within sectors to differences in mean latent ability, the earnings wedge τ^w_{jg} is equal to the mean latent ability in sector j of group g, and thus computing the change in average ability among researchers requires separating the two components of τ_{jg}.

Under this assumption, the model with gender and race estimates a 53.1% decline in ability, close to our baseline of 52.6%, whereas the model with immigrants estimates a decline in ability of 31.7%, as compared to the baseline number 37.6%, and thus infers a non-trivial increase in ability coming from an increase in the immigrant share of researchers. The intuition for this result is as follows: as shown in Figure 11, the trend in the share of researchers among immigrants is roughly parallel to the corresponding trend among non-immigrants. Absent differences in latent mean ability, this would predict that the difference in the mean log earnings of immigrant researchers from the mean log earnings of non-immigrant researchers should have stayed roughly constant. However, the data indicates that this difference has in fact expanded dramatically; assuming away group-specific returns to ability, this can only have happened because the mean latent ability among

Table A6: Group-specific targeted moments

moments:	employment shares						mean log earnings			
	S_g	S_{Cg}	$S_{R\|Cg}$	ΔS_g	ΔS_{Cg}	$\Delta S_{R\|Cg}$	E_{Rg}	E_{Hg}	ΔE_{Rg}	ΔE_{Hg}
By gender and race (1960-2021)										
white-male	0.609	0.107	0.103	-0.290	0.300	-0.032	0.460	0.368	0.196	0.182
white-female	0.271	0.094	0.026	0.006	0.373	0.006	0.434	0.336	0.196	0.159
other-male	0.081	0.043	0.096	0.130	0.247	-0.005	0.602	0.347	0.214	0.275
other-female	0.039	0.072	0.023	0.154	0.284	0.024	0.423	0.323	0.312	0.229
By birthplace (1970-2021)										
U.S.-born	0.944	0.133	0.075	-0.123	0.254	-0.027	0.619	0.474	0.067	0.038
foreign-born	0.056	0.143	0.171	0.123	0.249	-0.074	0.567	0.433	0.306	0.239

Notes: Group- and sector-specific employment shares and mean log earnings relative to group-specific mean log earnings of non-college graduates. Note that changes are taken between 2021 and 1960 for first 4 rows, whereas the last two rows take changes between 2021 and 1970.

Table A7: Estimated group-specific wedges

parameters	total wedge				earning wedge			
	τ_{Cg}	$\tilde{\tau}_{RHg}$	$\Delta\tau_{Cg}$	$\Delta\tilde{\tau}_{RHg}$	τ^w_{Rg}	τ^w_{Hg}	$\Delta\tau^w_{Rg}$	$\Delta\tau^w_{Hg}$
By gender and race (1960-2021)								
white-male	0	0	0	0	0	0	0	0
white-female	-0.086	-0.025	0.264	0.011	-0.166	-0.115	-0.418	0.282
other-male	-0.558	0.001	0.187	0.007	-0.336	-0.495	0.510	0.445
other-female	-0.255	-0.027	0.100	0.020	-0.325	-0.271	0.075	0.240
By birthplace (1970-2021)								
U.S.-born	0	0	0	0	0	0	0	0
foreign-born	0.053	0.020	-0.039	-0.005	0.040	0.014	0.638	0.133

Notes: To save table space, we define $\tilde{\tau}_{RHg} = \tau_{Rg} - \tau_{Hg}$ for Columns 2 and 4. Note that changes are taken between 2021 and 1960 for the first 4 rows, whereas the last two rows take changes between 2021 and 1970.

immigrants has increased dramatically relative to that of US-born workers.

Though the average change in researcher ability under this assumption seems plausible, the implied underlying differences in mean researcher ability are extremely implausible. For instance, attributing all cross-group differences to differences in mean ability implies an 85% immigrant ability premium in 2021, which is more than two standard deviations of the US-born ability distribution; in effect, the model suggests that an *average* immigrant's researcher ability is at the 95th percentile of the distribution of U.S.-born workers. Similarly extreme statements would result from the model with gender and race. We see two key takeaways from this exercise: first, that the estimated change in average researcher ability remains significant even allowing for differences in group-specific ability means, and second, that attributing all group-specific differences within sectors to differences in group-specific means results in implausible distributional differences; this exercise should thus be interpreted as at best a bounding exercise.

D.2.3 Group-specific dispersion of earnings

Figure A17 further confirms that, as predicted by our framework, the group-specific dispersion of earnings for researchers and other college graduates has been increasing over time. This is true within each group. This observation reconfirms that the composition effect from self-selection appears within each demographic group.

References

Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth.* Princeton, NJ:Princeton University Press.

Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange. 2012. "Changes in the Characteristics of American Youth: Implications for Adult Outcomes." *Journal of Labor Economics*, 30(4): 783–828.

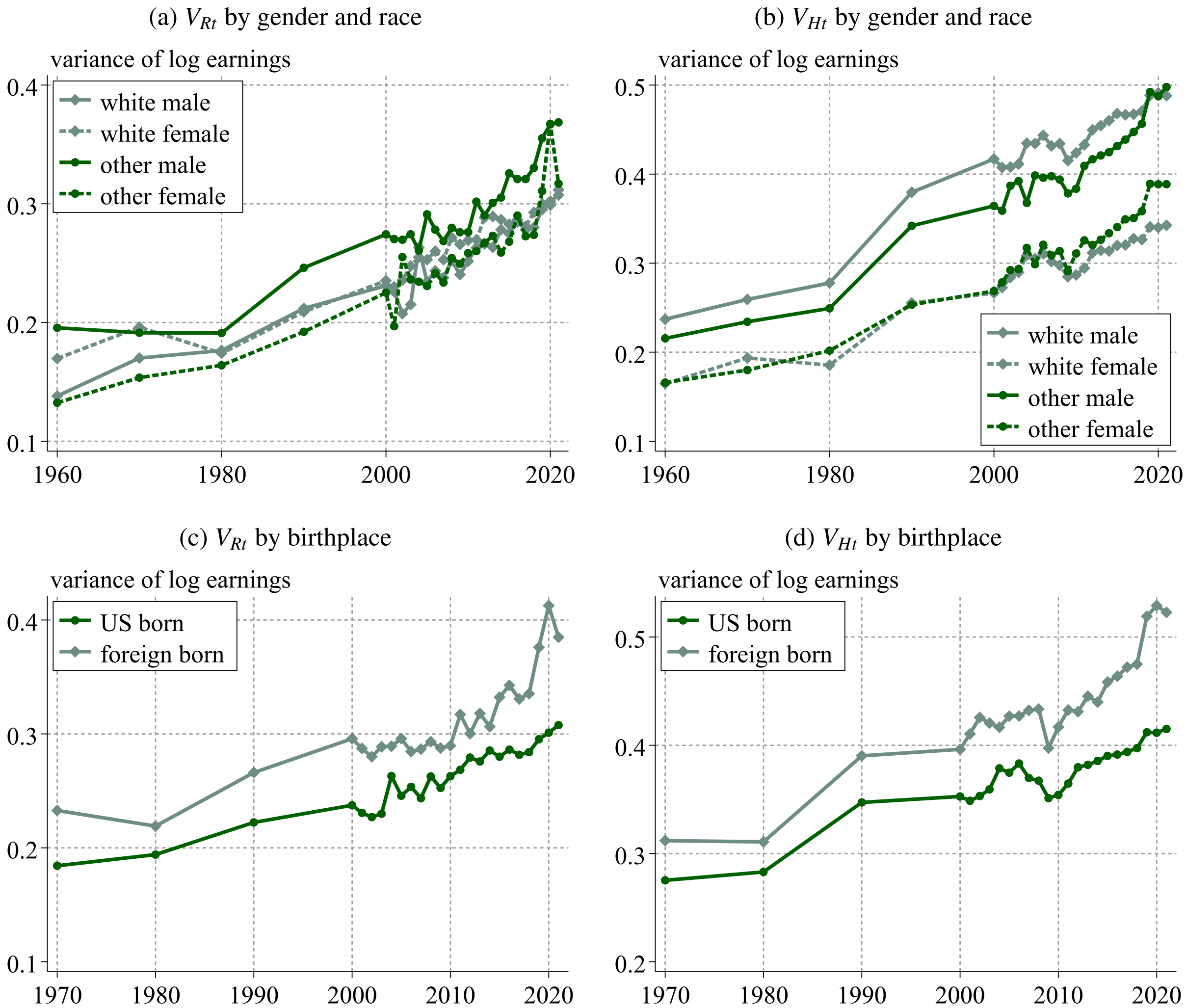
Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro. 2017. "Measuring the Sensitivity of Parameter Estimates to Estimation Moments." *Quarterly Journal of Economics*, 132(4): 1553–1592.

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics*, 104: 184–198.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. 2020. "Are Ideas Getting Harder to Find?" *American Economic Review*, 110(4): 1104–1144.

BLS. 2025. "National Longitudinal Survey of Youth 1979 cohort, 1979-2022 (rounds 1-30)." *U.S. Bureau of Labor Statistics (BLS). Produced and Managed by the Center for Human Resource*

Figure A17: Group-specific earnings dispersion



Research (CHRR), The Ohio State University.

BLS. 2026. "National Longitudinal Survey of Youth 1997 cohort, 1997-2023 (rounds 1-21)." *U.S. Bureau of Labor Statistics (BLS). Produced and Managed by the Center for Human Resource Research (CHRR), The Ohio State University*.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150–82.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Etienne Breton, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, David Van Riper, and Kari C.W. Williams. 2025. "IPUMS CPS: Version 13.0." *Minnesota Population Center, IPUMS*.

Jones, Benjamin F. 2009. "The Burden of Knowledge and the 'Death of the Renaissance Man':

Is Innovation Getting Harder?" *Review of Economic Studies*, 76(1): 283–317.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy. 2021. "Measuring Technological Innovation over the Long Run." *American Economic Review: Insights*, 3(3): 303–320.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018. "Life Cycle Wage Growth Across Countries." *Journal of Political Economy*, 126(2): 797–849.

NCSES. 2021. "National Survey of College Graduates, 1993-2021." *National Center for Science and Engineering Statistics (NCSES), National Science Foundation (NSF)*.

OECD. 2026*a*. "Economy: Productivity." *OECD Data Explorer, Organisation for Economic Co-operation and Development (OECD). Accessed in June 2020 through previous portal: https://stats.oecd.org*.

OECD. 2026*b*. "Science, Technology and Innovation: R&D." *OECD Data Explorer, Organisation for Economic Co-operation and Development (OECD). Accessed in June 2020 through previous portal: https://stats.oecd.org*.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rodgers, Jonathan Schroeder, and Kari C.W. Williams. 2025. "IPUMS USA: Version 16.0." *Minnesota Population Center, IPUMS*.

USPTO. 2025. "PatentsView: Baseline and Disambiguated Tables." *U.S. Patent and Trademark Office (USPTO), Department of Commerce*.