The Skill Premium in Times of Rapid Technological Change*

Tarek Hassan,[†] Aakash Kalyani,[‡] and Pascual Restrepo[§]

March 1, 2026

Abstract

This paper shows that the pace of technology creation is a key driver of the skill premium. It develops a model in which skilled workers have a comparative advantage in learning new technologies. As technologies age, they become standardized and accessible to other workers. The skill premium is determined by the interplay between the pace of technology creation and standardization. A rapid pace of technology creation leads to a sustained increase in the skill premium. We calibrate the model using novel text-based data on new technologies and their changing demand for skills as they age. These data show that new technologies are initially skill intensive but become less so as they age. The data also point to an increased pace of new technology creation starting in the 1970s and tapering off in the 2000s. In response to this rapid pace of technology creation, the model generates a 32 percent increase in the college premium, which begins to reverse in the 2010s. Our framework also explains why the college premium is higher in dense cities, why its increase was mainly urban, and why it rose first for young workers and later for older workers.

*The views expressed here are solely those of the authors and do not necessarily represent those of the Federal Reserve Bank of St Louis or the Federal Reserve System. We thank Erik Hurst, Loukas Karabarbounis, and Larry Schmidt for helpful comments.

[†]**Boston University**, 270 Bay State Rd, Boston, MA 02215; thassan@bu.edu.

[‡]**Federal Reserve Bank of St Louis**, Federal Reserve Bank Plaza, St. Louis, MO 63102; aakash.kalyani@stls.frb.org.

[§]**Yale University**, 28 Hillhouse Ave, New Haven, CT 06511; pascual.restrepo@yale.edu.

Introduction

The skill premium has risen sharply in the United States since the 1980s. Between 1965 and 1980, college-educated Americans earned about 1.5 times as much as those without a college degree. From 1980 to 2010, this premium doubled, with college graduates earning twice the wages of non-graduates. Since 2010, the college premium has flattened and begun a modest reversal. Figure 1 documents these well-known patterns using data from the Current Population Survey (CPS).[1]

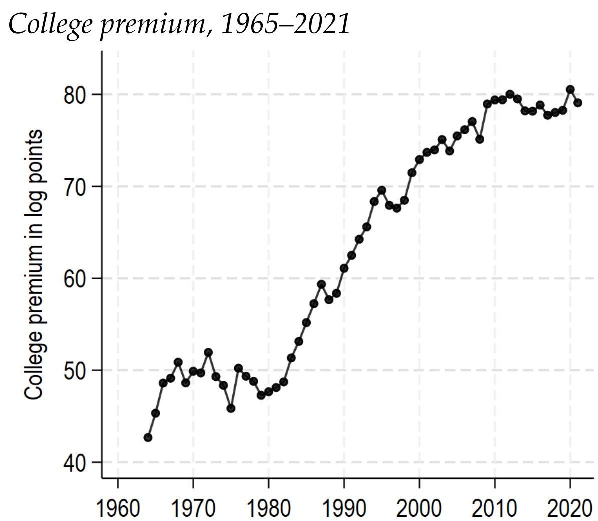

Figure 1: Skill premium in the US. The figure shows the wage college premium for the United States (from the Current Population Survey). The series is adjusted for demographics.

Existing theories attribute the rising skill premium to structural changes in production that favor skilled workers. Katz and Murphy (1992) emphasize skill-biased technical change—shifts in technology that raise the productivity of college workers relative to others. Krusell et al. (2000) highlight the rapid decline in equipment prices starting in the 1970s, which favors educated workers, who use equipment more intensively. Acemoglu and Restrepo (2022) argue that advances in automation reduced the demand for non-college workers by displacing them from tasks they performed. Others highlight changes brought by information technology (see Autor, Katz, and Krueger, 1998; Beaudry, Green, and Sand, 2016; Burstein, Morales, and Vogel, 2019).

This paper offers a complementary explanation, emphasizing the *pace of technology creation* as

[1]See Autor et al. (2008), Acemoglu and Autor (2011), and Autor (2019) for a summary of these trends.

a driver of the skill premium across time, locations, and worker cohorts. In our account, skilled workers are distinguished by their ability to learn to use new, recently invented technologies—a conception of skill that dates back to Schultz (1975). This perspective implies that the skill premium depends on the pace of technology creation: when this pace accelerates, the premium rises because skilled workers learn to use new technologies more rapidly. This mechanism differs from existing theories, which focus on whether successive waves of technology are intrinsically more skill-intensive or automated than earlier waves. Even if successive waves are identical in their characteristics, a faster pace of technology creation shifts the economy toward recently invented technologies, where skilled workers thrive.

This paper formalizes this mechanism and quantifies the contribution of changes in the pace of technology creation to the rise in the skill premium in the United States using novel text-based data on the creation and diffusion of new technologies.

We develop a model in which new technologies are introduced at an exogenous rate. Skilled workers learn to use them faster than less-skilled workers, but this advantage fades as technologies age, become standardized, and their use spreads. The skill premium depends on the pace of technology creation, the initial learning advantage of skilled workers, and how quickly this advantage diminishes. An increase in the pace of technology creation leads to a sustained increase in the skill premium, which reverses over time as technologies mature and less-skilled workers learn to use them.

The model can be quantified using data on the pace of new technology creation, their diffusion, and their changing demand for skills as they age. We use and extend the data from Kalyani et al. (2025), who combine patent text and job postings to identify new technologies and trace their diffusion in labor markets. We map skilled workers to those with a college education and use the model to interpret the evolution of the college premium.[2]

[2]Education is one observable proxy for the ability to learn new technologies, which is the notion of skill that matters in our theory. The presumption that educated workers have a comparative advantage in adapting to new technologies makes intuitive sense. As Schultz (1975) puts it, "The presumption is that education—-even primary schooling—enhances the ability of students to perceive new classes of problems, to clarify such problems, and to learn ways of solving them." Even if the ability to adapt and learn is innate rather than enhanced by education, one would expect quick learners to be highly educated on average. For previous work adopting this view see Greenwood and Yorukoglu (1997), Caselli (1999), Galor and Moav (2000), Mukoyama (2004), and Acemoglu and Restrepo (2018). For empirical support, see Bartel and Lichtenberg (1987) and Doms et al. (1997).

In line with *Schultz*'s view, the data show that new technologies initially demand more college-educated workers, but this reverses as technologies age. This evidence disciplines the magnitude of skilled workers' comparative advantage in new technologies and how fast it fades. The data also show a temporary increase in the pace of technology creation beginning in the 1970s, acceleration in the 1980s, and falling back in the 2000s.

In response to the rapid pace of technology creation during this period, the model generates a 28 log-point (32 percent) increase in the college premium between 1980 and 2010, which then flattens and begins to revert, precisely as in the data. Using the model, we decompose the evolution of the college premium into the contributions from changes in the supply of college-educated workers, changes in the pace of technology creation (our mechanism), and residual structural changes in the production process (as in Katz and Murphy, 1992). The total demand for college-educated workers increased by 100 log points since 1980. Changes in the pace of technology creation account for a third of the increase, while the remainder is attributed to residual structural changes in production, such as skill-biased technical change, capital deepening, and automation.

After presenting these results, we extend our theory to account for the diffusion of technologies across regions and from young to older workers. These extensions explain why the college premium was higher and rose the most in densely populated cities, and why it rose first for young workers and later for older workers.

The first extension explores the implications of technology diffusion across regions for the geography of wage inequality. As documented by Autor (2019), Rubinton (2020), and Eckert, Ganapati, and Walsh (2022), the college premium was higher and increased the most in dense cities. We extend the model to include locations with different population densities, in which new technologies diffuse from high- to low-density regions (as in Glaeser, 1999). Dense locations employ newer, more skill-intensive technologies, leading to a higher college premium and a faster increase than in the rest of the country.

We estimate the rate of technology diffusion across locations with varying population density using the data from Kalyani et al. (2025). The modal technology in the top 1% densest locations, such as New York and San Francisco, is 34 years old, while the modal technology in the bottom 50%

lowest-density locations is 48 years old, indicating sizable diffusion gaps. The estimated model matches these gaps and generates large differences in the steady-state level and growth of the college premium across regions. Our mechanism, coupled with the slow diffusion of technology, accounts for 6.2 of the 8.7 log-point differential increase in the skill premium between high- and low-density regions over 1980–2005.[3]

Our second extension explores the role of worker age. As shown by Card and Lemieux (2001), the college premium rose first among young workers, while older workers experienced a smaller, protracted increase. Card and Lemieux (2001) attribute this pattern to supply-side forces. Our extension shows that demand-side forces can also explain the age profile of the college premium. We introduce worker demographics and assume that younger workers have a comparative advantage in new technologies.[4] An increase in the pace of technology creation raises the skill premium among young workers using new technologies first. Only as technologies age and diffuse among older workers does the skill premium for them rise.

We calibrate this extension using Current Population Survey data on computer use by worker age. The data show that young workers initially used computers more intensively, with older workers catching up as the technology aged. In response to the rapid pace of technology creation from 1970 to 2000, the calibrated model generates age-specific increases in the college premium that account for half of the age gaps in the data.

Relationship to the literature: This paper contributes to the literature exploring how technology affects the skill premium, including work by Katz and Murphy (1992) and Krusell et al. (2000). Our paper identifies a novel driving force behind the rise in the skill premium: the rapid pace of technology creation from 1980 to 2000. We measure this pace directly, which allows us to quantify the contribution of this mechanism to changes in the skill premium. Our approach differs from Katz and Murphy (1992), who infer the extent of skill-biased technical change as a residual, but do not identify its drivers. Our micro-founded approach also allows us to explain the timing of the

[3]These findings complement Rubinton (2020) and Eckert et al. (2022), who propose mechanisms that amplify the effects of skill-biased technical change in urban areas.

[4]The idea that young workers are better at learning new technologies is consistent with evidence from cognitive and educational psychology (see Horn and Cattell, 1967; Salthouse, 1996; Lindenberger and Baltes, 1997; Burke and Barnes, 2006) and reduced-form evidence in economics (see Lehr, 2023).

rise in the skill premium, including its slowdown in 2010 and its predicted future trajectory.

Our approach is closer to Krusell et al. (2000), who identify capital prices as a driver of the skill premium. The difference with their work is that we emphasize college-educated workers' ability to use new technologies, while Krusell et al. (2000) emphasize their ability to use equipment, old or new. In our theory, the driving force is the pace of new technology creation, whereas in theirs it is the price of capital goods. Consider computers: in our theory, skilled workers initially use computers more frequently, but this difference diminishes as computers become easier to use. In their framework, skilled workers have a permanent advantage in using computers. This is why their theory predicts permanent effects from a decline in capital prices, while ours predicts long-lasting but temporary effects.[5]

The idea that workers differ in their ability to learn new technologies and the implications of this mechanism for inequality were explored in previous work by Greenwood and Yorukoglu (1997) and Caselli (1999).[6] Relative to these papers, our value added is in quantifying the contribution of this mechanism to wage inequality over time, space, and age groups, using novel text data. Greenwood and Yorukoglu (1997) propose a model in which skilled workers learn to use new technologies first, followed by other workers. They present comparative statics showing that technological breakthroughs temporarily raise inequality. Caselli (1999) emphasizes the role of heterogeneity in learning costs. When new technologies have high learning costs, their introduction brings inequality. When new technologies have low learning costs, their introduction reduces inequality. Caselli (1999) interprets the increase in inequality since the 1970s as being due to the arrival of a wave of technologies with high learning costs. This differs from our work, which emphasizes the effects of changes in the pace of new technology creation while keeping learning costs constant.

Finally, our work adds to a growing literature that traces new technologies and their diffusion patterns using textual analysis of patents and job postings: Kelly et al. (2021), Autor et al. (2023), Kogan et al. (2023), Kalyani (2024), and Kalyani et al. (2025). Our paper shows how these data can be used to study the drivers of inequality and the skill premium over time and across regions.

[5] Similar distinctions apply to the literature that focuses on the role of computer prices and IT, including Autor et al. (1998), Beaudry et al. (2016), and Burstein et al. (2019). Our model identifies the rapid pace of technology creation during this period as a key driver of the skill premium, beyond the specific nature of computers.

[6] Another branch focuses on how this affects technology adoption decisions (see Mukoyama, 2004).

Organization: Section 1 introduces our baseline model. Section 2 explains how we use text data to estimate the model and quantify the effects of changes in the pace of technology creation on the college premium. Section 3 provides quantitative results and decompositions. Section 4 explores the interplay of our mechanism with geography; Section 5 explores the interplay with worker age.

1 A Model of Technology, Standardization, and the Skill Premium

We first consider an economy in which only the pace of technology creation varies over time. We discuss the role of structural changes in the bias of technology and shifts in supply in Section 2.

New technologies arrive exogenously over time and are subsequently used to produce goods and services. Let $m(b)$ denote the mass of technologies introduced at date (cohort) b. At calendar time t, the mass of technologies of age u (i.e., born at $b = t - u$) is

$$m_t(u) \;=\; m\big(t-u\big).$$

Each of these technologies produces output according to

$$y_t(u) \;=\; A\big(t-u\big)\; z(u)\; \Big[\alpha(u)^{\frac{1}{\gamma}}\; h_t(u)^{\frac{\gamma-1}{\gamma}} \;+\; \big(1-\alpha(u)\big)^{\frac{1}{\gamma}}\; \ell_t(u)^{\frac{\gamma-1}{\gamma}}\Big]^{\frac{\gamma}{\gamma-1}},$$

where $h_t(u)$ and $\ell_t(u)$ denote the high-skill (college educated in the data) and low-skill (non-college) labor employed per technology of age u, respectively, and γ is their elasticity of substitution.

The terms $A\big(t-u\big)\; z(u)$ characterize technologies' productivity. The first term is cohort-specific and assumed to grow exponentially in birth date

$$A(b) \;=\; A_0\, e^{gb}, \text{ with } g > 0,$$

so that newer cohorts are more productive than previous ones, generating sustained growth.

The term $z(u)$ captures the productivity life-cycle of technologies. For example, technologies may mature and become more productive as they age. We assume

$$g \;>\; g_{\lim} = \lim_{u\to\infty} \frac{\dot{z}(u)}{z(u)},$$

so that old technologies eventually lose market share to new ones.

The function $\alpha(u)$ captures the skill intensity of technologies of age u. We assume, and show evidence, that $\alpha(u)$ is large for new technologies, but decreases with age u, so that skilled workers' comparative advantage at new technologies wanes over time. The level of $\alpha(u)$ represents this advantage; the slope measures how quickly the technology is standardized or how quickly knowledge of how to use it spreads.

Both $\alpha(u)$ and $z(u)$ are functions of technology age, but do not vary with calendar time. All technology waves are assumed to be identical, with the same productivity profile and skill intensity throughout their life cycle. This deliberate choice ensures that there are no structural changes in production shifting skill demand, and isolates the role of sheer changes in the pace of technology creation, summarized by the number of technologies in each cohort, $m(b)$.

Aggregate output at time t is a CES aggregate (with elasticity $\sigma > 1$) of all technologies:

$$Y_t \;=\; \left(\int_0^\infty m_t(u)\; y_t(u)^{\frac{\sigma-1}{\sigma}}\; du \right)^{\frac{\sigma}{\sigma-1}}.$$

The total supply of high-skill and low-skill labor is fixed at ℓ and h, respectively. Labor-market clearing requires that workers be employed across all technologies:

$$\int_0^\infty m_t(u)\, h_t(u)\, du \;=\; h, \qquad \int_0^\infty m_t(u)\, \ell_t(u)\, du \;=\; \ell.$$

Equilibrium: We treat the final good as the numeraire and normalize its price to 1. Given the path of technology vintages $\{m_t(u)\}$, an equilibrium is a sequence for output Y_t and real wages $\{W_{h,t},\, W_{\ell,t}\}$, so that markets clear. This yields three equilibrium conditions at each point in time:

i. A price-index condition for the final good,

$$1 \;=\; \int_0^\infty m_t(u) \left(\frac{c\big(\alpha(u),\, W_{h,t}, W_{\ell,t}\big)}{A(t-u)\, z(u)} \right)^{1-\sigma} du, \tag{1}$$

where

$$c\big(\alpha(u),\, W_{h,t}, W_{\ell,t}\big) \;=\; \Big[\, \alpha(u)\, W_{h,t}^{1-\gamma} \;+\; \big(1-\alpha(u)\big)\, W_{\ell,t}^{1-\gamma} \Big]^{\frac{1}{1-\gamma}}$$

is the unit cost of producing one unit of vintage u's output (with productivity normalized to 1). This cost depends on $\alpha(u)$—the weight placed on high-skill wages.

ii. The labor market for high-skill workers clears. This requires total wage payments to skilled labor, $W_{h,t}\, h$, to match the revenue it generates across all technologies:

$$W_{h,t}\, h \;=\; Y_t \int_0^\infty m_t(u) \left(\frac{c\big(\alpha(u), W_{h,t}, W_{\ell,t}\big)}{A\big(t-u\big)\, z(u)} \right)^{1-\sigma} \alpha(u) \left(\frac{W_{h,t}}{c\big(\alpha(u), W_{h,t}, W_{\ell,t}\big)} \right)^{1-\gamma} du. \tag{2}$$

iii. The labor market for low-skill workers clears:

$$W_{\ell,t}\, \ell \;=\; Y_t \int_0^\infty m_t(u) \left(\frac{c\big(\alpha(u), W_{h,t}, W_{\ell,t}\big)}{A\big(t-u\big)\, z(u)} \right)^{1-\sigma} \big[1-\alpha(u)\big] \left(\frac{W_{\ell,t}}{c\big(\alpha(u), W_{h,t}, W_{\ell,t}\big)} \right)^{1-\gamma} du. \tag{3}$$

The three equations determine unique equilibrium values for the endogenous variables $Y_t, W_{h,t}$, and $W_{\ell,t}$ at each point in time. These depend on the pace of technology creation, summarized by $m_t(u)$, which is the driving force in the model.

Our first proposition shows that, when the pace of technology creation is constant, the model admits a balanced growth path (BGP).

Proposition 1 (Balanced Growth Path)**.** *Suppose technology creation is constant, i.e. $m(b) = m$, so that $m_t(u) = m$. There exists a unique BGP along which real wages and output grow at rate g, the skill premium remains constant and is independent of m.*

All proofs are in the appendix. To illustrate the proposition, consider the case with $\gamma = \sigma$.

When $m_t(u) = m$, the economy aggregates to a CES production function in h and ℓ:

$$Y_t = A(t)\, m^{\frac{1}{\sigma-1}} \Big(\underbrace{\Big(\int_0^\infty e^{-(\sigma-1)gu}\, z(u)^{\sigma-1}\, \alpha(u)\, du\Big)^{\frac{1}{\sigma}}}_{\equiv \alpha_h} h^{\frac{\sigma-1}{\sigma}} + \underbrace{\Big(\int_0^\infty e^{-(\sigma-1)gu}\, z(u)^{\sigma-1}\, [1-\alpha(u)]\, du\Big)^{\frac{1}{\sigma}}}_{\equiv \alpha_\ell} \ell^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma}{\sigma-1}}.$$

The weights α_h and α_ℓ give the overall skill intensity of the economy and determine the skill premium along the BGP:

$$\frac{W_{h,t}}{W_{\ell,t}} = \left(\frac{\alpha_h}{\alpha_\ell}\right)^{\frac{1}{\sigma}} \left(\frac{h}{\ell}\right)^{-\frac{1}{\sigma}}. \tag{4}$$

This expression is reminiscent of the model in Katz and Murphy (1992). The difference is that the weights α_h and α_ℓ are micro-founded and depend on the skill requirements and productivity life cycles of existing technologies.

This special case illustrates several properties of the balanced growth path:

- Quality improvements across cohorts, subsumed in $A(t)$, drive growth over time.
- Along the BGP, old and new technologies coexist. Technologies of age u account for constant shares of employment and output. The life cycle of technology's market share depends on the growth rate of $z(u)$—how fast incumbent technologies improve—relative to g—competition from new entrants. Old technologies eventually lose market share at a rate $(\sigma - 1)(g - \lim_{u\to\infty} \frac{\dot{z}(u)}{z(u)}) > 0$.
- The aggregate skill intensity of the economy is a weighted average of the skill intensity of technologies of age u, $\alpha(u)$. The weights $e^{-(\sigma-1)gu}\, z(u)^{\sigma-1}$ capture market shares.
- The BGP skill premium depends on how fast technologies are standardized (governed by $\alpha(u)$) relative to how fast they lose market shares as they age (governed by $z(u)$).
- The BGP skill premium is independent of m. Increases in m shift the *level* of wages and GDP but do not affect the skill premium in the long run. This is because m scales the mass of

technologies across all ages proportionally, while leaving their long-run age distribution—the key determinant of the skill premium—unchanged.

- As usual, the ratio h/ℓ of labor endowments reduces the skill premium.

Changes in the pace of technology creation. The next proposition shows that an increase in the pace of technology creation temporarily increases the skill premium.

Proposition 2 (Changes in the Pace of Technology Creation)**.** *Consider an economy on its balanced growth path at time t_0. An increase in $m(b)$ at t_0 from m to $m' > m$, whether permanent or temporary, generates a transitory increase in the skill premium.*

For intuition, again consider the case $\gamma = \sigma$, in which the economy aggregates to

$$Y_t \;=\; A(t)\,\Big(\underbrace{\big(\int_0^\infty m_t(u)\, e^{-(\sigma-1)gu}\; z(u)^{\sigma-1}\; \alpha(u)\; du\big)}_{\equiv \alpha_{h,t}}^{\frac{1}{\sigma}}\; h^{\frac{\sigma-1}{\sigma}}$$
$$+ \underbrace{\big(\int_0^\infty m_t(u)\, e^{-(\sigma-1)gu}\; z(u)^{\sigma-1}\; [1-\alpha(u)]\; du\big)}_{\equiv \alpha_{\ell,t}}^{\frac{1}{\sigma}}\; \ell^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\sigma}{\sigma-1}}.$$

The weights $\alpha_{h,t}$ and $\alpha_{\ell,t}$ are now time varying. This variation is driven by changes in the pace of technology creation, summarized by $m_t(u)$, which determines the distribution of technologies by age. The skill premium along the transition is now

$$\frac{W_{h,t}}{W_{\ell,t}} \;=\; \left(\frac{\alpha_{h,t}}{\alpha_{\ell,t}}\right)^{\frac{1}{\sigma}}\;\left(\frac{h}{\ell}\right)^{-\frac{1}{\sigma}}. \tag{5}$$

An increase in $m(b)$ starting at t_0 shifts the distribution of technology by age toward newer technologies in the *short run*. Because skilled workers have a comparative advantage in these technologies (i.e., $\alpha(u)$ is decreasing in u), $\alpha_{h,t}$ rises relative to $\alpha_{\ell,t}$, raising the contribution of high-skill labor to output, and consequently the skill premium.

The increase in the skill premium is transitory. Over time, the distribution of technology by age returns to its initial steady-state level. This is true for both temporary and permanent increases in $m(b)$, since in the latter case, new technologies eventually age. This aligns with Proposition 1, which shows that the long-run skill premium is independent of m.

Propositions 1 and 2 clarify the link between the pace of technology creation, the age distribution of technology, and the skill premium. In the long run, the economy's skill bias depends on how quickly technology is standardized ($\alpha(u)$) and how rapidly older technologies fall behind ($z(u)$). In the short run, a rapid pace of technology creation, $m(b)$, places greater weight on newer, more skill-intensive technologies, raising the skill premium.

We now explore these implications in the data.

2 Data and Measurement

Our quantitative exercise assumes that the economy is on its balanced growth path (BGP) in 1970. It then experiences an acceleration in the pace of technology creation, $m(b)$, over the period 1980–2000, which generates a long-lasting increase in the skill premium. Our objective is to quantify the magnitude of this increase. Doing so requires measuring the rate of arrival of new technologies $m(b)$, the skill bias and standardization of technologies of different ages $\alpha(u)$, and their productivity life-cycle, $z(u)$.

This section describes the data used, highlighting how we identify the birth of new technologies from patent text data and track their diffusion in job postings.

2.1 Measuring the emergence of new technology and its demand for skills

We build on Kalyani et al. (2025), who trace the origin and diffusion of new technologies from 1976 to 2007. The basic idea is to treat Wikipedia technology pages as distinct technologies in our model. These technologies are then traced across patents and job postings through the technical bigrams commonly used to refer to them. We follow their data construction procedure, with minor changes to extend their data back in time.

In a first step, we isolate technical bigrams—two-word combinations that appear frequently in US patents but are not part of common English usage. We compare the full text of all utility patents filed between 1930 and 2012 at the US Patent and Trademark Office with a historical corpus of English usage (COHA) covering the period from 1810 to 1930 (see Davies, 2010). We retain bigrams that (i) do not appear in the historical corpus and (ii) occur in at least 100 patents over the sample period.

In a second step, we link technical bigrams to distinct technologies in Wikipedia. A bigram is linked if, when entered into the Wikipedia search engine: (i) at least one of the first five search results links to a page describing a technology; and (ii) that page contains the bigram in its title, in the summary, or at least ten times in the body of the entry. Following Kalyani et al. (2025), we define a Wikipedia page as describing a technology based on the presence of application- and device-oriented section titles, and exclude pages focused on problems, management, or business concepts.[7]

This yields a set of Wikipedia technology pages, each linked to multiple technical bigrams. Each Wikipedia page represents a distinct technology, and the bigrams linked to it are the expressions commonly used in patent texts to describe it. For example, the bigrams "cable fiber" and "optic communication" are linked to the Wikipedia page for Optical fiber, which we treat as a single technology in our dataset.

In a third step, we measure the emergence year of each technology in our sample. We do so by counting the number of patent mentions of bigrams associated with the technology. We define the emergence year as the first year in which (i) this mention count reaches at least 100 and (ii) the number of mentions doubles over the subsequent five years, indicating a surge in research activity related to the technology.[8]

This procedure yields a list of 6,259 technologies with emergence years between 1941 and 2006. Each technology is associated with a unique Wikipedia page and linked to technical bigrams that can be traced across patent text and job postings. Table 1 lists the top technologies in each year

[7]Specifically, section titles must contain at least one of the words "application(s)," "use(s)," "type(s)," "operation," "characteristic(s)," "feature(s)," "device(s)," "technical," and "commercial" and none of the words "responses," "mitigation," "problems," "causes," "signs," "symptoms," "adverse effects," "management," "manager," "risk assessment," "business model," "distribution model," "customer," "strategy," and "service provider."

[8]Our patent data cover the period 1930–2012. Because our definition of emergence requires both a minimum number of mentions and subsequent growth over a five-year window, we treat the 1930–1940 period as a burn-in and exclude the final five years of the sample, during which post-emergence dynamics are not observed.

since 1941, defined by the most associated patenting activity, together with their most mentioned bigram. These technologies include *pulse generator* in the 1950s, *polymerase chain reaction* in the 1990s, and *tablet computers* in 2006.

Table 1: Top technologies by year of emergence

Title	**Top Bigram**	**Year**	**Title**	**Top Bigram**	**Year**
motor drive	drive_motor	1941	graphics processing units	processing_unit	1974
recording head	recording_head	1942	tert-butyl alcohol	tert_butyl	1975
spring-loaded camming device	spring_loaded	1943	immortalised cell line	cell_line	1976
voltage divider	output_voltage	1944	lateral flow test	reaction_zone	1977
adipic acid	adipic_acid	1945	growth factor	growth_factor	1978
corrosion inhibitor	corrosion_inhibitor	1946	nucleic acid	nucleic_acid	1979
dichloromethane	methylene_chloride	1947	targeted drug delivery	drug_delivery	1980
programmable interval timer	output_signal	1948	solar cell	solar_cell	1981
chemical reaction	reaction_condition	1949	bovine serum albumin	bovine_serum	1982
methacrylic acid	methacrylic_acid	1950	operational data store	data_stored	1983
pulse generator	pulse_generator	1951	holographic optical element	optical_element	1984
polymer solution	polymer_solution	1952	personal computer	personal_computer	1985
level shifter	voltage_level	1953	communications system	communication_system	1986
electrical element	circuit_element	1954	user interface	user_interface	1987
pull-up resistor	resistor_connected	1955	polymerase chain reaction	chain_reaction	1988
electrically conductive adhesive	electrically_conductive	1956	polyethylene glycol	polyethylene_glycol	1989
epoxy	epoxy_resin	1957	universal remote	remote_control	1990
propylene glycol	propylene_glycol	1958	electrical connector	electrically_connected	1991
chain-growth polymerization	polymer_chain	1959	internet protocol suite	internet_protocol	1992
flowchart	flow_chart	1960	cable & wireless communications	wireles_communication	1993
ac power plugs and sockets	ac_power	1961	carboxymethyl cellulose	carboxymethyl_cellulose	1994
pulse-width modulation	pulse_width	1962	real-time polymerase chain reaction	pcr_product	1995
antioxidant	free_radical	1963	omega-3 acid ethyl esters	ethyl_ester	1996
laser beam machining	laser_beam	1964	sense strand	sense_strand	1997
duty cycle	duty_cycle	1965	mobile device	mobile_device	1998
phase-shift keying	phase_shift	1966	carbon nanotube	carbon_nanotube	1999
particle-size distribution	particle_size	1967	marker-assisted selection	desired_trait	2000
printed circuit board	printed_circuit	1968	snp array	polymorphism_snp	2001
methyl acetate	methyl_ester	1969	silo	bulk_storage	2002
multiplexer	input_signal	1970	wi-fi 6	wi_fi	2003
pressure measurement	pressure_sensor	1971	coaxial cable	coaxial_cable	2004
compacted oxide layer glaze	oxide_layer	1972	linoleic acid	linoleic_acid	2005
size-exclusion chromatography	pore_size	1973	tablet computers	tablet_computer	2006

Notes: The table reports the technology that emerged in each year with the most patent mentions. The table also reports the most mentioned bigram for each technology.

In a fourth step, we trace the diffusion of these technologies in the labor market using Lightcast job postings data from 2010 to 2023. Lightcast collects postings from online job boards and employer websites and provides information on job location, detailed occupational codes, and the full text of each posting. We flag a posting as involving a given technology if its text contains an associated bigram. As shown in Kalyani et al. (2025), 90% of mentions refer to tasks that involve using the technology or support its operation.

In the final step, we measure the skill demands of each technology using job posting data. Following Kalyani et al. (2025), we infer skill requirements from the job's detailed 6-digit occupation, which provides information on the education needed for the position. We assign each job posting the average share of college-educated workers in its target occupation, computed from the 2010 American Community Survey. For example, a job posting for a judge is assigned a college intensity of 98 percent, while a posting for a paralegal is assigned 45 percent.

The resulting dataset spans 300 million job postings from 2010–2023, allowing us to observe how technologies from various cohorts—from those that emerged in the 1940s to those that emerged in the 2000s—change their skill demands and employment as they age.

2.2 The pace of technology creation $m(b)$ during 1941-2006.

Figure 2 plots the empirical counterpart of $m(b)$, computed as the number of technologies that emerged in each year. The dashed line depicts the raw data. The solid line reports a smoothed version, computed by fitting a local polynomial. We focus on the smoothed version, which isolates low-frequency changes in the pace of technology creation that affect the equilibrium skill premium in our model.

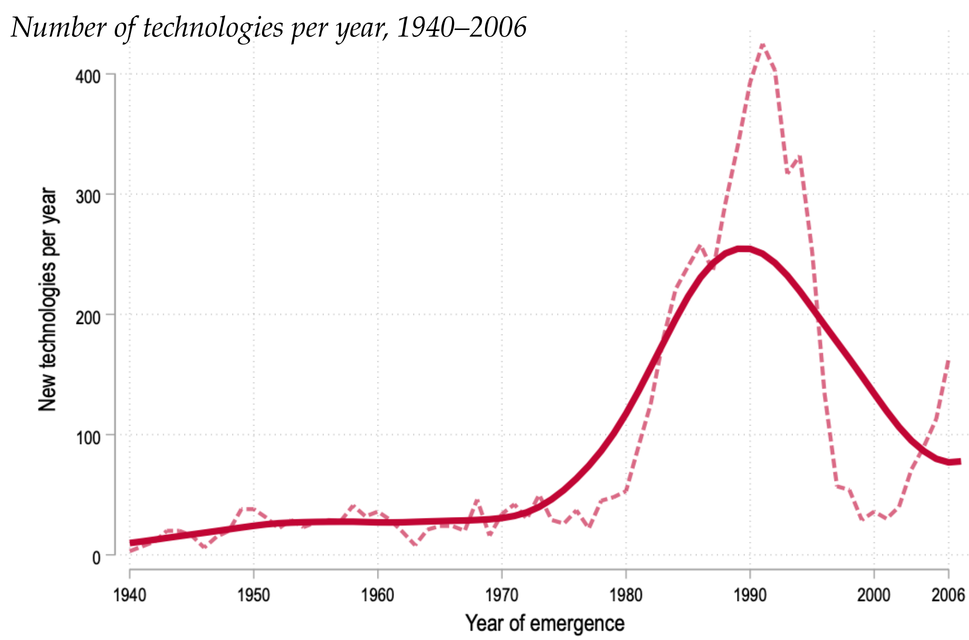

Figure 2: Time series for the pace of new technology creation, $m(b)$. The figure plots the number of new technologies created each year. The dashed line shows the raw data, and the solid line is a smoothed version that isolates low-frequency variation.

Prior to 1970, the pace of technology creation was stable, with 25–30 new technologies introduced each year. This pace rose sharply during the 1980s, reaching a peak of 250 technologies per year by the late 1980s. The data also show a subsequent slowdown, with the pace of technology creation declining to 100 technologies per year during 2000–2007.

By design, the series for $m(b)$ captures a specific dimension of technological change: the introduction of distinct technologies at the extensive margin. In our model, it is this form of technological churn—the arrival of new, differentiated technologies that workers must learn to use—that affects the skill premium. The series is not, and should not be viewed as, a measure of total technological progress in the US economy. The measure excludes improvements at the intensive margin (such as performance gains or cost reductions in existing technologies and their diffusion), advances in managerial practices and quality control, and other disembodied productivity improvements that raise TFP without generating identifiable new technologies.

Examining the technologies that emerged between 1980 and 1995 helps illustrate the types of innovations driving the increase. While the list spans a wide range of domains, three broad strands are prominent. One strand is an ICT wave, progressing from early hardware advances (e.g., magnetic storage) to later software and internet-related technologies (e.g., internet protocols and user interfaces). A second strand comprises semiconductor and fabrication innovations that enabled the production of smaller, more powerful digital components with applications across multiple industries. A third strand is a boom in molecular biology, driven by advances in our understanding of DNA, leading to technologies such as targeted drug delivery, gene therapy, and genetically modified organisms.

The pattern shown in Figure 2 aligns with independent evidence of an increased pace of extensive-margin technology creation during the 1980s and 1990s. For example, TFP growth data from Fernald (2014) exhibits a secular decline since the 1950s, interrupted by a boom in the late 1990s and early 2000s. The timing of this boom is precisely what one would expect from an *earlier* burst of technology creation in the 1980s and 1990s, given that the new technologies underlying the increase in $m(b)$ take time to reach full impact.

The period from 1980 to 1995 also saw a large increase in patent applications, driven in part by expanded innovation opportunities in biotechnology, information technology, and software industries—sectors that also feature prominently in our data—, increased funding by VC investors, and a shift toward more applied research (see Kortum and Lerner, 1999).[9] In line with this interpretation, this period featured large investments in new digital technologies, alongside elevated startup formation (see Decker et al., 2016).

We conduct a wide range of robustness checks designed to address four concerns: the representativeness of Wikipedia, changes in language over time, inflation in patenting activity, and the possibility that the pattern is driven solely by ICT. Figure 3 summarizes the robustness series, normalized relative to their 1960 level:

Number of technologies per year, 1940–2006

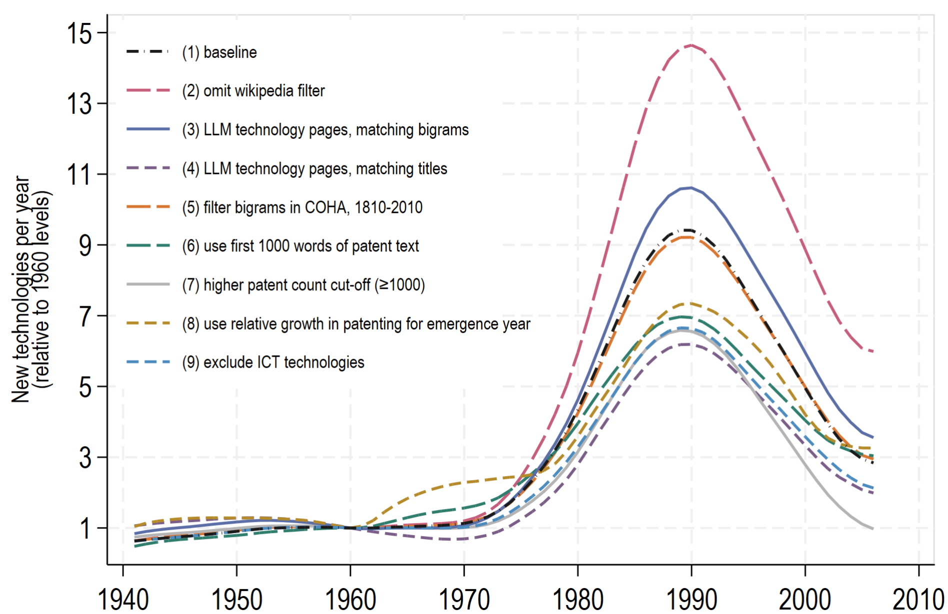

Figure 3: Robustness checks: time series for the pace of new technology creation, $m(b)$**.** This figure plots alternative measures of the number of new technologies created each year. All series are smoothed using local polynomial regression and normalized to 1 in 1976 for comparability.

- *Series (2)* removes the Wikipedia filter and plots the raw number of bigrams that emerged each year (using the same definition of emergence introduced above). This robustness check

[9] A different hypothesis is that the rise in patenting was due to the Federal Court Improvements Act of 1982. As shown by Kortum and Lerner (1999), US firms also increased patenting during this period *outside* of the US, where these legal changes had no impact. This supports the view that the early rise in patents over 1980–1995 was due to an increase in US innovation.

shows that, even when we set aside Wikipedia pages, the number of technical bigrams in patents increases in the 1980s and 1990s. The increase is more pronounced, suggesting that the Wikipedia filter mitigates the proliferation of overlapping technical terms, but does not alter the timing of the rise.

- *Series (3)* refines our list of Wikipedia technologies by using a Large Language Model (Claude Haiku 4.5) to drop pages that do not reference actual technologies. We then trace the emergence year of these technologies as above, by linking them to patents through common technical bigrams. This curated set of technologies exhibits a slightly larger increase, starting in 1970 and peaking in 1990.

- *Series (4)* goes one step further and dispenses entirely with the use of technical bigrams to identify or trace the emergence of technologies. The series starts from the same refined set of technologies as Series (3), identified via a Large Language Model, but dates their emergence by directly matching these technologies to patents based on their titles. The resulting trends in the pace of technology creation remain similar, though the peak is less pronounced.

- *Series (5)* further explores the robustness of the first step described above by also removing contemporary English language from patents to isolate technical language. This is done by excluding words that appear in a corpus of the common English language (COHA) from 1810 to 2010. This check shows that the pattern in Figure 2 is not due to changes over time in language.

- *Series (6)* reconstructs our series for $m(b)$ using only the first 1,000 words in a patent. This check shows that the pattern in Figure 2 is not an artifact of rising patent length over time.[10]

- *Series (7)* increases the patent-mention threshold from 100 to 1,000, requiring technical bigrams to appear in ten times as many patents to be included in the analysis. This reduces the number of bigrams and technologies in the sample, but the resulting trends in the pace of technology creation remain similar, though slightly less pronounced. This exercise is

[10] The timing of the increase in patent length is also inconsistent with this interpretation. Median patent length rose slowly from 2,500 words in 1940 to 3,000 in 1980 and 3,200 in 1990. The vast increase in patent lengths came later, between 1995 and 2015, with median patent length doubling from 3,500 to 7,000 words.

conservative, since many technologies from the 1980-1990 boom have not had sufficient time to reach this higher mention count.

- *Series (8)* adjusts our construction for the growth in patenting over time, giving more weight to early mentions and less weight to mentions in later years. Following Lerner and Seru (2022), we weight a mention in year t by the ratio of the sample's average number of patents per year to the number of patents filed in t. This series times the year of emergence of a technology based on how fast its patenting grew as a share of the overall flow of patents. The series shows a sizable increase, starting in 1960, accelerating in the 1980s, and peaking in 1990. While this check is conservative (the rise in patenting in the 1980s and 1990s seems to reflect increased innovation), the results show that the pattern in Figure 2 is not an artifact of higher mention counts in recent years driven by the overall growth in patenting.

- *Series(9)* removes information and communication technologies (ICT), defined as those whose technical bigrams are mentioned in patents classified under G06, H04, H01, and H03 (i.e., patent classes related to information and communication technologies). This series shows that the rapid pace of technology creation during the 1980s and 1990s was not exclusive to ICT but reflected a broader increase in other fields.

All series indicate a stable pace of technology creation from 1940 to 1960, followed by a sharp increase during the 1980s and 1990s. Across specifications, the number of technologies introduced per year increases by a factor of 7 to 10 over this period (with the exception of Series 2, which rises by a factor of 14, presumably because of duplication), then declines.

2.3 Estimating $\alpha(u)$ and $z(u)$

We set the elasticities γ and σ externally and estimate the functions $\alpha(u)$ and $z(u)$ using the job-posting data. For the elasticities of substitution, we set $\gamma = 1.4$, as in Katz and Murphy (1992). We also set the elasticity of substitution between technologies to $\sigma = 3.5$ to align with the median estimate of elasticities of substitution across varieties in Broda and Weinstein (2006).[11] Finally, we

[11]The aggregate elasticity of substitution between college and non-college workers in our model is higher than γ, due to substitution across technologies. This effect is quantitatively minor, which justifies calibrating γ to match the

set $g = 2\%$ per year to match the long-run growth rate of the US economy.

Estimating $\alpha(u)$: We estimate $\alpha(u)$ to match the changing demand for college workers in a given technology as it ages. As documented in Kalyani et al. (2025), new technologies predominantly generate employment opportunities for high-skill workers. Over time, as technologies mature, they increase their demand for non-college workers. Panel A of Figure 4 replicates this finding using our expanded dataset. The figure plots the average college intensity of job postings associated with technologies of age u, ranging from $u = 4$ (the first time we observe them in the Lightcast data in 2010) to $u = 82$.

Panel A: Skill intensity $\frac{h(u)}{h(u)+\ell(u)}$ *of technology* *Panel B: Relative demand* $\log(\frac{h(u)}{\ell(u)}) + \gamma \ln(\frac{W_h}{W_\ell})$

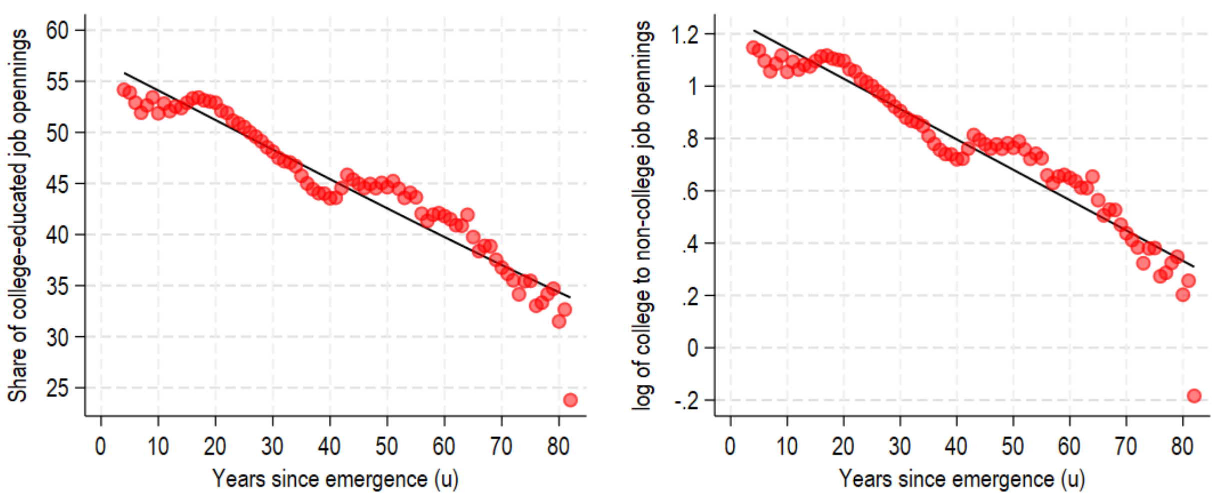

Figure 4: Reduced skill intensity of technologies as they age. Panel A plots the average skill intensity of job postings by technology age, computed as the share of college-educated workers $\frac{h(u)}{h(u)+\ell(u)}$ for technologies of age u. Panel B plots the relative demand for college labor, computed as $\log(\frac{h(u)}{\ell(u)}) + \gamma \ln(\frac{W_h}{W_\ell})$, where $\gamma = 1.4$ and wages are averaged over 2010–2023. Red markers represent data. Solid black lines show the fitted values from estimating equation (6).

The figure shows a clear decline in the skill intensity of technologies as they age. In their year of emergence, 57 percent of jobs involving a new technology require a college degree, compared with 34 percent of jobs involving technologies that are 80 years old. This pattern supports the view that skilled workers have a comparative advantage in new technologies that fades as technologies

aggregate elasticity of substitution in Katz and Murphy (1992).

age, become standardized, and knowledge about their use diffuses.

To capture this pattern, we parameterize $\alpha(u)$ using a logistic function:

$$\alpha(u) \; = \; \frac{\exp(\theta_0 - \theta_1 u)}{1 + \exp(\theta_0 - \theta_1 u)}.$$

Here, θ_0 gives the initial advantage of college labor at new technologies and θ_1 the rate at which the advantage diminishes (either because of standardization or knowledge diffusion).[12] The relative demand for college workers by technology age is

$$\frac{h(u)}{\ell(u)} = \frac{\alpha(u)}{1 - \alpha(u)} \left(\frac{W_h}{W_\ell}\right)^{-\gamma}.$$

Taking logs and rearranging yields the estimating equation

$$\ln\left(\frac{h(u)}{\ell(u)}\right) + \gamma \ln\left(\frac{\bar{W}_h}{\bar{W}_\ell}\right) = \theta_0 - \theta_1 u + \epsilon_u, \tag{6}$$

where the error term is interpreted as a measurement error. The left side is computed from the data on skill demand by technology age, setting $\gamma = 1.4$, and using the average college premium $\bar{W}_h/\bar{W}_\ell$ over 2010–2021 (to adjust for the role of prices during the period we observe job postings).

Estimating (6) by OLS yields $\theta_0 = 1.26$ (s.e. $= 0.020$) and $\theta_1 = 0.012$ (s.e. $= 0.001$). The functional form fits the data well, with an R^2 of 90.8%. The solid black lines in both panels show the fitted relationship and the model-implied college intensity by technology age.

Table A1 in the Appendix reports robustness checks for our estimates of $\alpha(u)$. The table reports estimates obtained by excluding ICT technologies from the sample, as well as estimates for the early (2010–2015) and late (2016–2023) years in our job-posting sample. The estimates are similar to our baseline, suggesting that the decline in skill demand as technologies age is not specific to ICT technologies and is stable across periods.

The table also reports estimates of (6) obtained by aggregating the data to technology-cohort (b) and calendar-year (t) cells. This panel allows us to trace the evolution of skill demand within

[12]The data cannot separate the role of standardization and the increasing availability of knowledge on how to use a technology. For our purposes, both forces are isomorphic and play identical roles.

each technology cohort over the 2010–2023 period, accounting for calendar-year or cohort effects (using dummies for birth cohorts from 1940–1960, 1960–1980, 1980-2000, and 2020 onward).[13]

Year fixed effects absorb aggregate labor market conditions and changes in coverage in the Lightcast data. Controlling for these effects, we estimate a standardization rate of $\theta_1 = 0.012$. Cohort fixed effects, by contrast, absorb differences in the skill intensity of past technology waves. Controlling for cohort effects yields a higher estimated standardization rate of $\theta_1 = 0.022$. This shows that the pattern shown in Figure 4 is present within technology cohorts and is not driven by newer vintages being more skill-intensive than older ones.

Estimating $z(u)$: The function $z(u)$ controls the path of productivity as a technology ages. We estimate $z(u)$ to match the share of workers using technologies of age u, from the job-posting data.

Panel A of Figure 5 plots the average share of job postings (regardless of education level) that mention technologies of age u. The figure reveals a clear life-cycle pattern: Employment shares are initially small, at 0.4% for young technologies, then rise steadily and peak at 2% 35 years after emergence. That is, it takes, on average, 35 years for a technology to be fully deployed in the US labor market. Beyond that point, employment shares decline gradually with age.[14]

The above patterns are informative of the behavior of $z(u)$. In the model, the share of all workers employed per technology of age u (i.e., $(h(u) + \ell(u))/(h + \ell)$) is

$$\text{Employment share}(u) \;=\; \text{constant}\; e^{-(\sigma-1)gu}\; z(u)^{\sigma-1}\; \underbrace{\frac{\alpha(u)W_h^{-\gamma} + (1-\alpha(u))W_\ell^{-\gamma}}{c(\alpha(u), W_h, W_\ell)^{\sigma-\gamma}}}_{\equiv \kappa(\alpha(u), W_h, W_\ell)}. \tag{7}$$

The term $\kappa(\alpha(u), W_h, W_\ell)$ adjusts for differences in technology costs due to their skill mix. These depend on $\alpha(u)$ and the wage rates faced by firms. The term $e^{-(\sigma-1)gu}$ captures the erosion of incumbents' market shares due to competition by new frontier technologies. The term $z(u)^{\sigma-1}$ links productivity, $z(u)$, to labor market shares.

[13] In this panel specification, one could estimate γ directly by placing the college wage premium term on the right side and exploiting its changes over time. Doing so yields point estimates ranging from 1.6 to 2.2, depending on the covariates, which are in the ballpark of the standard literature.

[14] These shares are computed as a fraction of all job postings associated with technologies 82 years old or younger, as these are the ones in our sample.

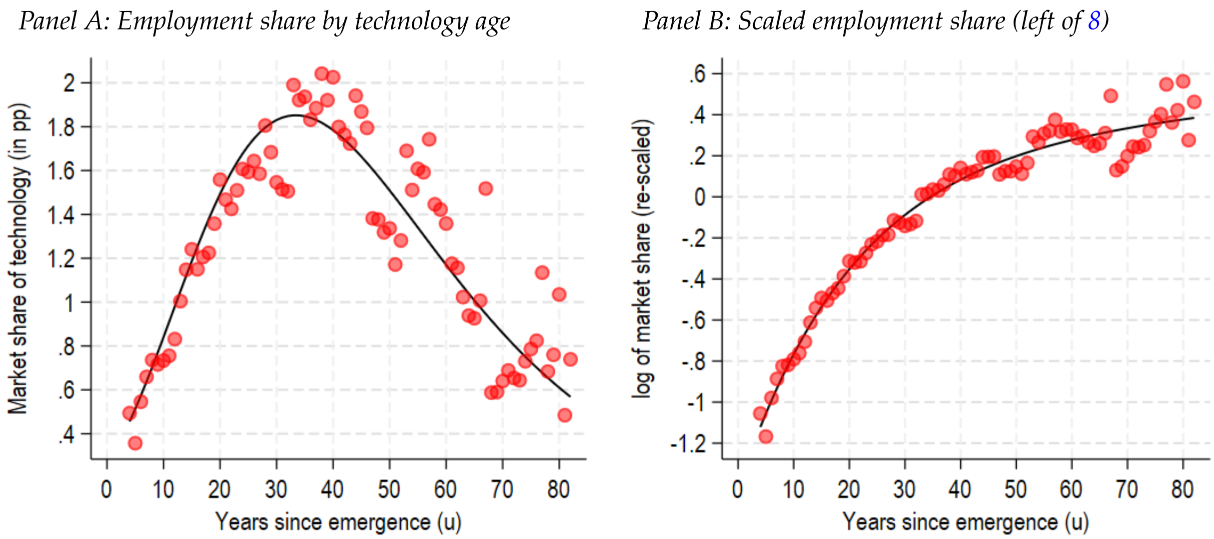

Figure 5: Life cycle of technologies' employment shares. Panel A plots the share of employment per technology of age u (including both college and non-college workers). This is reported as a fraction relative to technologies under 82 years old. Panel B plots the left side of the estimating equation (8), $\frac{1}{\sigma-1}\ln \text{Employment share}(u) + gu - \frac{1}{\sigma-1}\ln \kappa(\alpha(u), W_h, W_\ell)$. Red markers show the data. Solid black lines show the fitted values from the non-linear least squares estimation of (8).

To match the inverse U-shape pattern shown in Panel A of Figure 5, we parameterize $z(u)$ as

$$\ln z(u) = g_m u + \frac{1}{\lambda}(g_M - g_m)\,(1 - e^{-\lambda u}).$$

This specification normalizes $z(0)$ to 1 and allows its growth rate to change with age:

$$\frac{\dot{z}(u)}{z(u)} = g_m + (g_M - g_m)e^{-\lambda u}.$$

g_M gives $z(u)$'s initial growth rate. This switches at a rate λ and converges to g_m. The specification generates the observed market-share dynamics if $g_M > g$—new technologies first gain market share—and $g > g_m$—old technologies lose market share to new ones.

Taking logs in (7), rearranging, and plugging the functional form for $z(u)$ yields the equation

$$\frac{1}{\sigma-1}\ln \text{Employment share}(u) + gu - \frac{1}{\sigma-1}\ln \kappa(\alpha(u), \bar{W}_h, \bar{W}_\ell) = \frac{1}{\sigma-1}\ln \text{constant} + g_m u + \frac{1}{\lambda}(g_M - g_m)e^{-\lambda u} + \varepsilon_u, \qquad (8)$$

where ε_u is a measurement error term. This equation shows that $z(u)$ can be recovered from labor market shares, adjusting for the growth of frontier technologies (the $g\,u$ term) and differences in cost due to their skill intensity (the κ term).[15]

The left side of (8) can be measured using job posting data, along with average market wages from 2010–2023 and our previous estimates of $\alpha(u)$, γ, and σ, to construct the adjustment terms. Panel B of Figure 5 plots the left side of (8) against technology age u. The resulting curve reveals the shape of $z(u)$: it increases rapidly at early ages, identifying g_M, and then decelerates as technologies mature, identifying λ and g_m.

Fitting (8) via non-linear least squares yields $g_m = 0.2\%$ (se= 0.003), $g_M = 8.2\%$ (se= 0.010), and $\lambda = 0.048$ (se= 0.009). The model fits the data well, with an R^2 of 97.2%. The fitted values are shown by the solid line in Figure 5. As intended, the model matches the life-cycle of employment shares by technology age.

Table A2 in the Appendix reports a series of robustness checks, including estimates that exclude ICT technologies or use the early and late years of the job-posting data, to assess the stability of the estimates. The table also reports estimates of equation (8) by aggregating the data to technology cohort (b) and calendar year (t) cells. As before, we control for year effects, which capture national labor market trends and changes in coverage in the Lightcast data, and cohort effects. Across all specifications, the estimates for $z(u)$ generate a similar inverse U-shape for employment shares as our baseline. This shows that the pattern above, where technologies initially gain market share and are then slowly pushed out by new entrants, is a robust finding and is not driven by differences across cohorts, ICT technologies, or other national trends.

Matching the BGP college premium: The BGP level of the college premium is pinned down by $\alpha(u)$, $z(u)$, and the relative supply of college-educated workers, h/ℓ. In our baseline exercise, we fix h/ℓ at its average post-2010 CPS value of 0.53. Doing so aligns with the period during which we observe job-posting data, ensuring that the skill demands shown in Figure 4 and the skill supply

[15] This parallels a large literature in industrial organization and trade that measures productivity by inverting market shares (Syverson, 2004; Foster et al., 2008; De Loecker, 2011). Note that z is only identified up to a multiplicative constant. We normalized $z(0) = 1$. This is without loss of generality since multiplying $z(u)$ by a constant scales all aggregates by the same amount and leaves the college premium unchanged.

are measured in the same units. We later relax this choice and study the effects of changes in the supply of college-educated workers.

To match the BGP level of the skill premium in 1970, we rely on the asymptotic behavior of $z(u)$. While the data identify $z(u)$ for technologies up to 82 years old, the rate at which older technologies lose market share is not observed. This limit obsolescence rate is given by $g - \lim_{u\to\infty} \frac{\dot{z}(u)}{z(u)} > 0$. The higher this rate, the lower the share of old technologies used in the labor market, and the higher the BGP skill premium. Motivated by this logic, we assume a constant obsolescence rate beyond age 80 and set it to 0.35% per year to match a BGP college premium of 50 log points in 1970.[16]

3 The Skill Premium Over Time.

We next use the estimated model and our measured series for $m(b)$ to quantify how the rapid pace of technology creation in the 1980s and 1990s affected the evolution of the US college wage premium.

3.1 Changes in the pace of technology creation and the skill premium

We assume the economy was in its BGP before 1970, reflecting the stable annual rate of new-technology creation during this period, shown in Figure 2. We then feed in the changes in the pace of technology creation measured over 1970–2007. We use the smoothed series for $m(b)$ shown by the solid line in Figure 2 to isolate its low-frequency variation. We keep all other determinants of the college premium fixed.

Figure 6 reports the model-implied effects on the college premium (black) and compares them to the data (red). In response to the rapid pace of technology creation during 1970–2000, the model generates a 28 log-point increase in the college premium, closely matching both the timing and magnitude observed in the data. The model also captures the flattening of the college wage premium post-2000, as technologies introduced in the 1970s and 1980s became standardized and widely used by less-skilled workers.

[16] We also explored an alternative approach that matches the initial college premium by adjusting the limit behavior of $\alpha(u)$ for old technologies. While both the asymptotic behavior of $\alpha(u)$ and $z(u)$ matter for the level of the skill premium, the choice of which one is used to match the initial BGP level has little effect on the estimated changes in inequality.

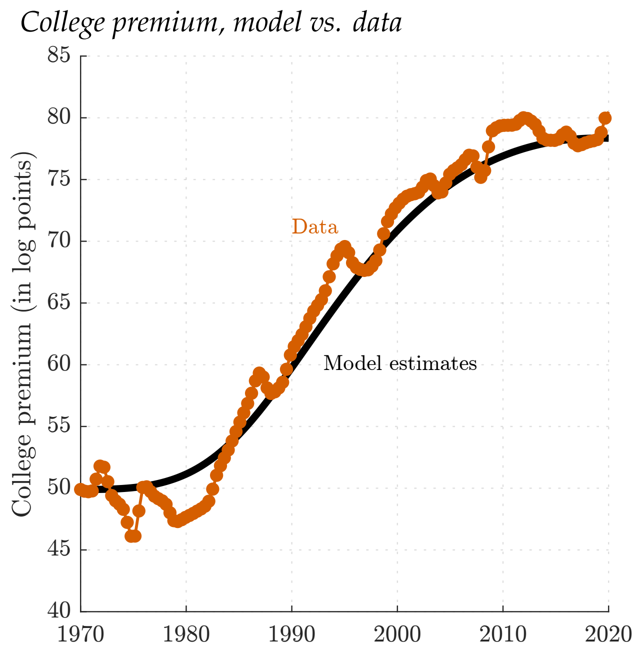

Figure 6: Skill premium over time: Model vs. data. The figure compares the college wage premium from the model (black line) with the data (red line). The model series shows the predicted path of the college premium in response to the rapid pace of technology creation $m(b)$ between 1970 and 2007, holding all other factors constant. The data series is computed from the Current Population Survey and shows the log wage differential between college and non-college workers. Both series are measured in log points.

The estimated shapes of $z(u)$ and $\alpha(u)$, recovered directly from the textual data, explain why the model's predictions match the observed behavior of the college premium. The $z(u)$ profile implies that technologies reach peak economic relevance 35 years after introduction, while the $\alpha(u)$ profile indicates slow standardization. Together, these features imply that the temporary acceleration in technology creation beginning in the 1970s generates a protracted, decades-long, increase in the college premium that gradually tapers off, consistent with the data.

The model also delivers predictions for the future path of the skill premium. Figure 7 reports the model-implied college premium from 1970 to 2080, computed under the assumption that there are no further changes in the pace of technology creation, $m(b)$, after 2007. In this scenario, the college premium remains elevated, but decreases to 68 log points by 2080. This shows that even temporary increases in the pace of technology creation can generate high levels of inequality that persist for generations.

Figure 8 reports the model's implications for wages, output, and TFP. Panel A shows the

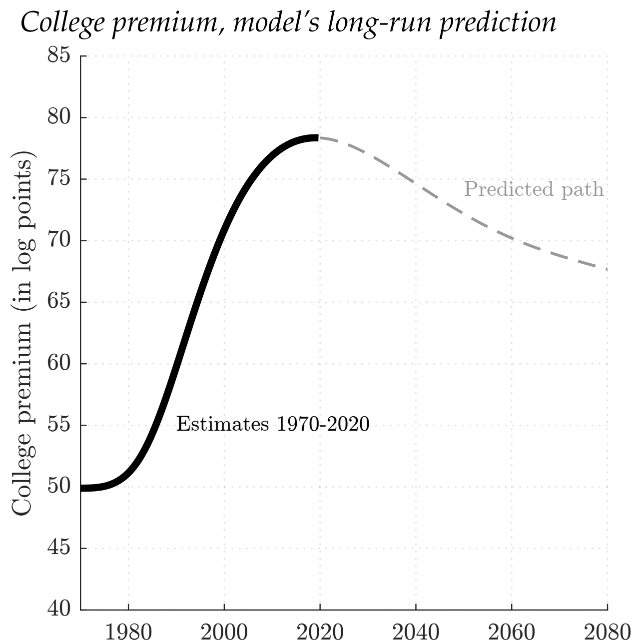

Figure 7: Predicted behavior of the skill premium in the future. The figure plots the model-predicted path for the college wage premium, assuming no further changes in $m(b)$ after 2007. All other model parameters are held constant.

effects of changes in the pace of technology creation, $m(b)$, on wage and output levels. The acceleration in technology creation beginning in the 1970s initially raised the wages of college-educated workers, starting in the early 1980s. Only around 1990 did these technologies become sufficiently standardized to raise wages for non-college workers.

Panel B reports the implications of the model for annual total factor productivity (TFP) growth, also computed under the assumption that there are no further changes in $m(b)$ after 2007. Our model generates a boom-bust cycle that qualitatively aligns with US trends. The rapid pace of technology creation starting in the 1970s led to a lagged increase in annual TFP growth, which peaked in 1995 and rose from 2% to 2.4% per year. TFP growth then reverts to its initial BGP level with some minor overshooting around 2040.

This dip in productivity growth results from the fact that new technology cohorts in the 2020s are small relative to the large mass of incumbents, reducing their incremental impact on output. The model cannot generate the persistent slowdown in the long-run TFP growth rate of the economy observed in the US data, since the long-run growth rate of output, g, is taken as

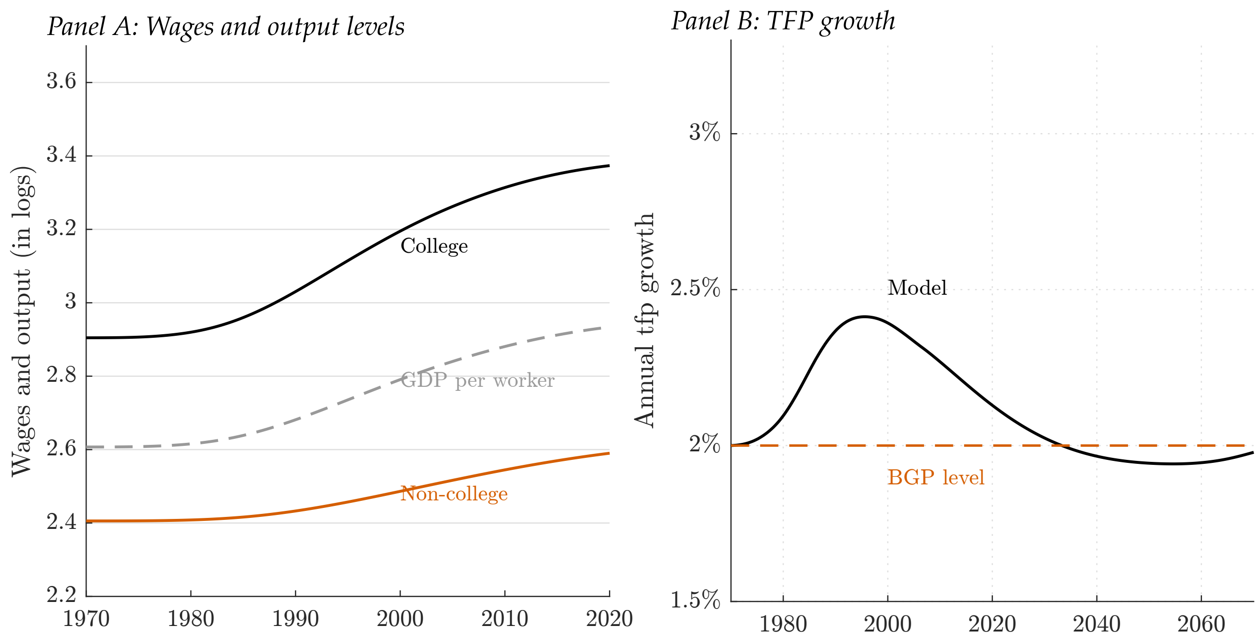

Figure 8: Implications for wage levels, output, and productivity growth. The figure shows the predicted path for wages, output, and TFP in response to the rapid pace of technology creation $m(b)$ between 1970 and 2007, holding all other factors constant. Panel A plots the log deviations from trend for GDP (dashed line), wages of college-educated workers (orange line), and wages of non-college workers (blue line). Panel B plots annual TFP growth rates.

exogenous and held constant in these exercises.

3.2 Decomposing changes in the college premium

The previous section assumed the supply of skilled workers and the skill bias of technology were fixed in (calendar) time. We now account for changes in these determinants of the skill premium.

The supply of college and non-college workers, h_t and ℓ_t, are computed using the Current Population Survey. The relative supply, $\frac{h_t}{\ell_t}$, rose from .15 in 1970 to .6 in 2020.

To allow for structural changes in production that increase the demand for college-educated workers across technologies—beyond those driven by technology age—we augment our specification of α with time-varying shifters θ_t:

$$\alpha_t(u) = \frac{\exp(\theta_t - \theta_1 u)}{1 + \exp(\theta_t - \theta_1 u)}.$$

As in Katz and Murphy (1992), these shifters capture, in reduced form, the roles of capital prices, automation, and other demand-side drivers of the college premium.

We recover θ_t as a residual, computed so that the calibrated model matches the college-premium data when we feed in changes in (i) the supply of college-educated labor h_t/ℓ_t, (ii) the residual shifters θ_t, and (iii) the pace of technology creation $m(b)$.

Panel A of Figure 9 reports the recovered series for θ_t (black). The series indicates that θ_t increased at an average rate of 2.1% per year. For comparison, the red series reports estimates of θ_t that do not account for changes in the pace of technology creation, $m(b)$. This series attributes the entire increase in demand for college workers to structural changes in production, as in Katz and Murphy (1992). In the absence of our mechanism, matching the data would require a larger residual, growing at 4.0% per year.

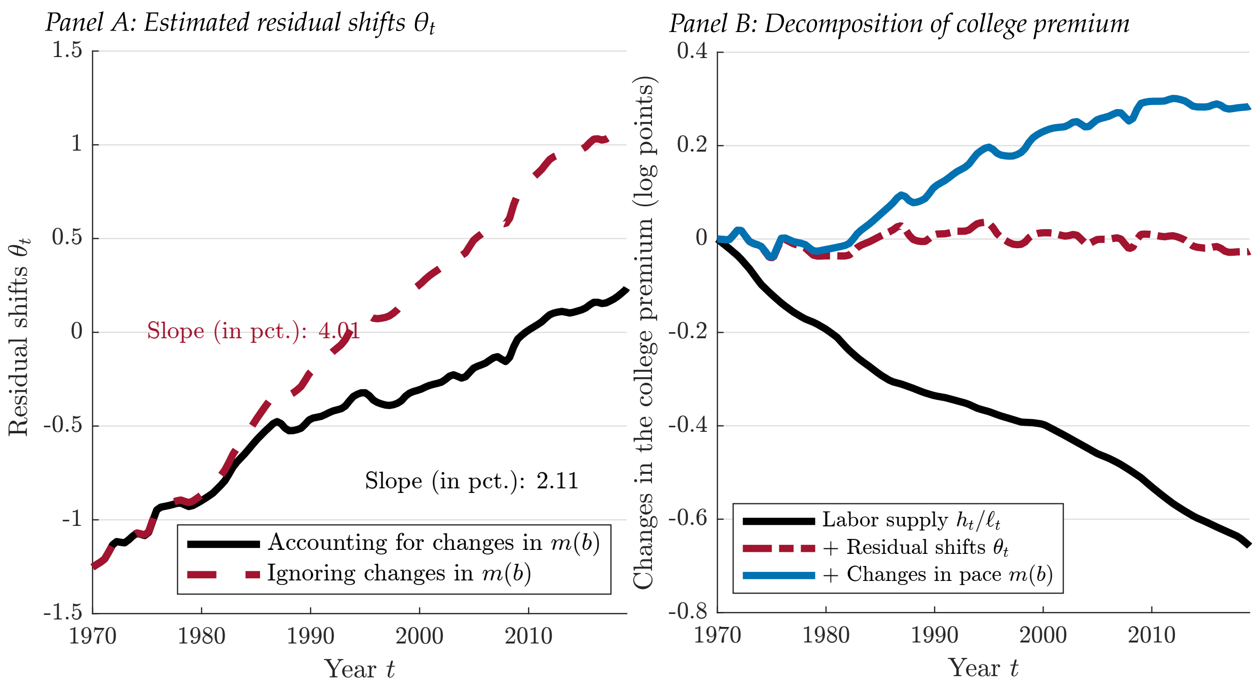

Figure 9: Decomposition of college premium. Panel A plots estimated residual shifts in technology θ_t, accounting for changes in $m(b)$ (black line) and without accounting for changes in $m(b)$ (red line). Panel B decomposes observed changes in the college premium since 1970 into: (1) the effect of changes in the relative supply of college workers h_t/ℓ_t (black line); (2) residual shifts in technology θ_t (red line); and (3) changes in the pace of technology creation $m(b)$ (blue line).

Panel B of Figure 9 decomposes the observed changes in the college premium into contributions from changes in supply, changes in the pace of technology creation, and residual changes in the skill bias of technology. The black line shows the contribution of supply changes. The large increase in the number of college-educated workers in the US would, on its own, have reduced the college premium by 70 log points. The 100 log-point gap between this counterfactual and the data reflects the total shift in relative demand for college-educated workers over this period.

The red dotted line captures the contribution of residual shifts in technology, which raise the college premium by 70 log points. The blue line incorporates the effects of changes in the pace of technology creation, $m(b)$. By construction, this line matches the data. In this decomposition, the rapid pace of technology creation during the 1980s and 1990s explains a third of the 100 log-point increase in the relative demand for college-educated workers, with residual forms of skill-biased technical change accounting for two-thirds.

The conclusion that changes in the pace of technology creation account for a third of the total shift in demand for college-educated workers is robust across different parameter values. Most notably, Table A5 in the Appendix summarizes the decomposition results for different values of γ—the elasticity that mediates the magnitude of supply-side forces. In scenarios varying γ from 1.2 to 2, changes in the pace of technology creation explain 34.4% to 25.3% of the total change in demand for college-educated workers since 1970.

4 The Skill Premium across Regions

This section explores the implications of spatial diffusion of technology for the geography of inequality. Figure 10 plots the college premium across commuting zones of different densities, using Census and American Community Survey (ACS) data. The pattern mirrors the findings of Autor (2019), Rubinton (2020), and Eckert et al. (2022). The college premium was already higher in dense areas by 1980 and rose more rapidly there in subsequent decades, increasing by 33.5 log points in dense cities versus 24.7 log points in less dense areas by 2007.[17]

We show that slow diffusion of new technologies from dense to less dense areas can account for

[17] We focus on 1980 as the starting year because the 1970 Census provides limited geographic coverage for low-population areas. In that year, data are reported only for county groups with at least 250,000 residents, which makes it difficult to infer outcomes in low-density commuting zones. This limitation is less severe in subsequent years, which report data for county groups or Public Use Microdata Areas (PUMAs) with populations of 100,000 or more. Even for these years, empirical patterns in low-density regions should be interpreted with caution, as the available data do not allow for precise measurement. We also focus here on men, as the trends for women across regions are noisier.

College premium across regions, 1980–2007

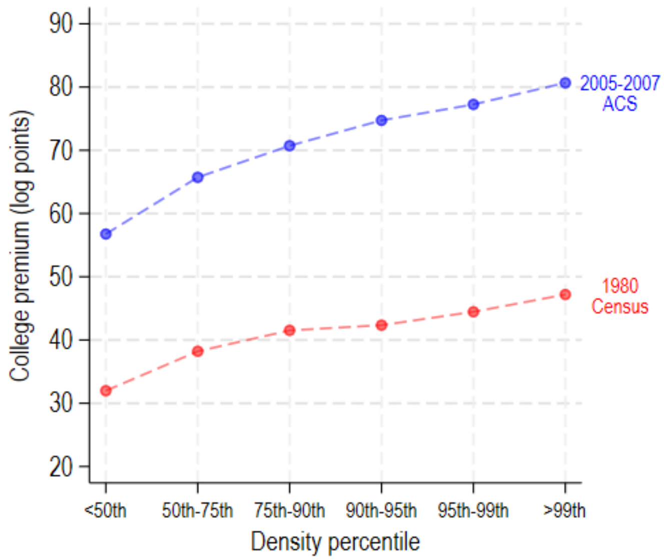

Figure 10: The skill premium across regions. The figure plots the college wage premium by population density. The dashed lines plot data for the 1980 Census (in red) and the 2005–2007 American Community Survey (in blue). Commuting zones are grouped in fixed density bins, according to their population density in the 2005–2007 American Community Survey (ACS).

both the higher level of the college premium in urban locations and its more pronounced increase.

Modal age of technology across regions

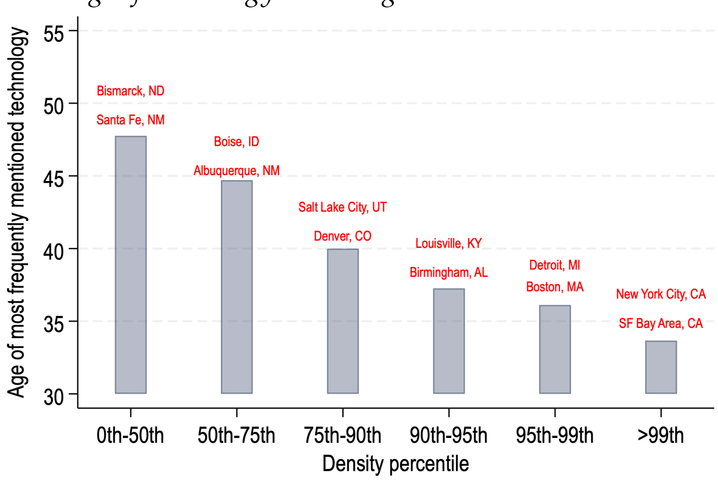

Figure 11: Modal technology age across commuting zones by density. The figure reports modal ages (i.e., the one with the highest employment share) of technologies used in US commuting zones by density bins, based on their 2005–2007 population density.

The idea that technologies diffuse slowly across space is supported by our data. Figure 11 plots the modal age of technologies across US regions of varying density. In the bottom 50% lowest-density regions, the modal technology is 48 years old. By contrast, in the top 1% highest-density regions, the modal technology is 34 years old, indicating sizable diffusion lags.

4.1 A model with spatial diffusion

Consider an economy with multiple locations, indexed by their density d. We abstract from trade and migration, and assume locations only interact via technology diffusion.

As before, $m(b)$ denotes the mass of technologies introduced at b and $m_t(u) = m\big(t-u\big)$ the mass of age u technologies at calendar time t. Once created, a technology of age u diffuses to a location of density d with probability $p(u,d)$. THis function is increasing and log-submodular in $\langle u, d\rangle$, implying that technology arrives first in high-density areas, while low-density regions catch up over time, consistent with Figure 11.

Once a technology of age u diffuses to location d, it produces output

$$y_t(u,d) \;=\; A\big(t-u\big)\; z(u)\left[\alpha(u)^{\frac{1}{\gamma}}\; h_t(u,d)^{\frac{\gamma-1}{\gamma}} \;+\; \big(1-\alpha(u)\big)^{\frac{1}{\gamma}}\; \ell_t(u,d)^{\frac{\gamma-1}{\gamma}}\right]^{\frac{\gamma}{\gamma-1}},$$

where $h_t(u,d)$ and $\ell_t(u,d)$ are labor inputs, and $A(b) \;=\; A_0\, e^{gb}$ with $g > \lim_{u\to\infty} \dot{z}(u)/z(u)$.

A location's aggregate output at time t is a CES aggregate of all technologies:

$$Y_t(d) \;=\; \Big(\int_0^{\infty} p(u,d)\; m_t(u)\; y_t(u,d)^{\frac{\sigma-1}{\sigma}}\; du\Big)^{\frac{\sigma}{\sigma-1}}.$$

Labor-market clearing requires all workers in each location to be employed:

$$\int_0^{\infty} p(u,d)\; m_t(u)\, h_t(u,d)\, du \;=\; h(d), \qquad \int_0^{\infty} p(u,d)\; m_t(u)\, \ell_t(u,d)\, du \;=\; \ell(d).$$

Here $h(d)$ and $\ell(d)$ are exogenous labor endowments by region.

Locations differ only in the technologies they use, with these differences captured by the diffusion probability $p(u,d)$. There are no other regional differences affecting the skill bias of production, and we again keep the relative supply $h(d)/\ell(d)$ constant across regions to emphasize the role of technology diffusion.

Equilibrium and effects of changes in the pace of technology creation: Given a path of technology creation $\{m_t(u)\}$, an equilibrium involves sequences for output $Y_t(d)$ and real wages $\{W_{h,t}(d), W_{\ell,t}(d)\}$ that vary over time and across regions due to changes in technology creation and diffusion. The equilibrium conditions are in the Appendix and parallel those in Section 1.

Proposition 3 (Balanced Growth Path across Regions)**.** *Suppose the pace of technology creation is constant, i.e.* $m(b) = m$*, so that* $m_t(u) = m$*. There exists a unique balanced growth path along which real wages and output grow at rate* g*, the skill premium in locations of density* d *is constant in time, independent of* m*, and increases in* d*.*

The proposition shows that the model admits a balanced growth path where the skill premium rises with population density. This is because high-density regions operate newer, more skill-intensive technologies, whereas low-density regions operate older, more standardized technologies.

Proposition 4 (Changes in the Pace of Technology Creation and Effects across Space)**.** *Assume* $\sigma = \gamma$ *and* $\alpha(u)$ *is sufficiently high for recently introduced technologies. Consider an economy in its balanced-growth path at time* t_0*. An increase in* $m(b)$ *at* t_0 *from* m *to* $m' > m$*, whether permanent or temporary, generates a transitory increase in the skill premium at all locations. The increase is more front-loaded in high-density locations: for any* $d' > d$*, the skill premium rises more in* d' *than in* d *early on in the transition.*

The proposition shows that an increase in the pace of technology creation raises the skill premium, with a stronger initial effect in high-density regions. This is because new technologies reach these regions first and only later diffuse to low-density areas.

The requirement that $\sigma = \gamma$ ensures that the elasticity of substitution between high-skill and low-skill labor does not vary across regions. Otherwise, differences in this elasticity affect how much relative wages move. The requirement that $\alpha(u)$ is sufficiently high for new technologies ensures that skilled workers in dense regions are more exposed to new technologies than low-skill workers. Our quantification below explores what happens more generally when $\sigma \neq \gamma$ and $\alpha(u)$ is estimated from the data.

Propositions 3 and 4 clarify the link between the pace of technology creation, spatial diffusion, and changes in the skill premium across regions. We now turn to the data to examine these implications and to quantify the contribution of this mechanism to the regional trends in the college premium shown in Figure 10.

4.2 Estimating the diffusion process across space $p(u, d)$

To explore the quantitative implications of Propositions 3 and 4, we extend our estimation procedure to account for differences in technology diffusion across space.

As above, we set $\gamma = 1.4$, $\sigma = 3.5$ and $g = 2\%$ per year. We keep the estimation of $\alpha(u)$ unchanged since the patterns shown in Figure 4 are identical within high and low-density regions (see Table A4 in the appendix). We also calibrate the limit obsolescence rate to match a BGP college premium of 46 log points seen in the data for cities at the 90th–95th percentiles of density. All variation in the college premium across regions and over time is left completely untargeted.

The main difference is that, in the regional model, the average employment share per technology at age u varies across locations due to the diffusion lags $p(u, d)$. Panel A of Figure 12 plots the share of employment (regardless of education level) per technology of age u, computed from the job posting data. The figure shows these data separately for high-density (in blue, the top 1% densest commuting zones) and lower-density regions (in red, commuting zones between the 25th and 30th density percentiles). The plot confirms that the distribution of technology use in dense regions is shifted towards younger technologies relative to that in lower-density regions.[18]

We jointly estimate $p(u, d)$ and $z(u)$ to match the data on employment shares by density. The employment share per technology of age u (i.e., $(h(u, d) + \ell(u, d))/(h(d) + \ell(d))$) is

$$\text{Employment share}(u, d) = \text{constant}(d)\, p(u, d)\, e^{-(\sigma-1)gu}\, z(u)^{\sigma-1}\, \kappa(\alpha(u), W_h(d), W_\ell(d)). \quad (9)$$

The first term on the right is a location-specific constant. The second term captures the arrival of technology. The remaining terms are the same as in (7), with the term $\kappa(\alpha(u), W_h(d), W_\ell(d))$

[18]The shares are computed as a fraction of all job postings associated with technologies 80 years old or younger, as these are the ones in our sample. Note also that the figure does not condition on technologies arriving at a place.

Panel A: Employment share(u) per technology.

Panel B: Re-scaled employment share (left of 10).

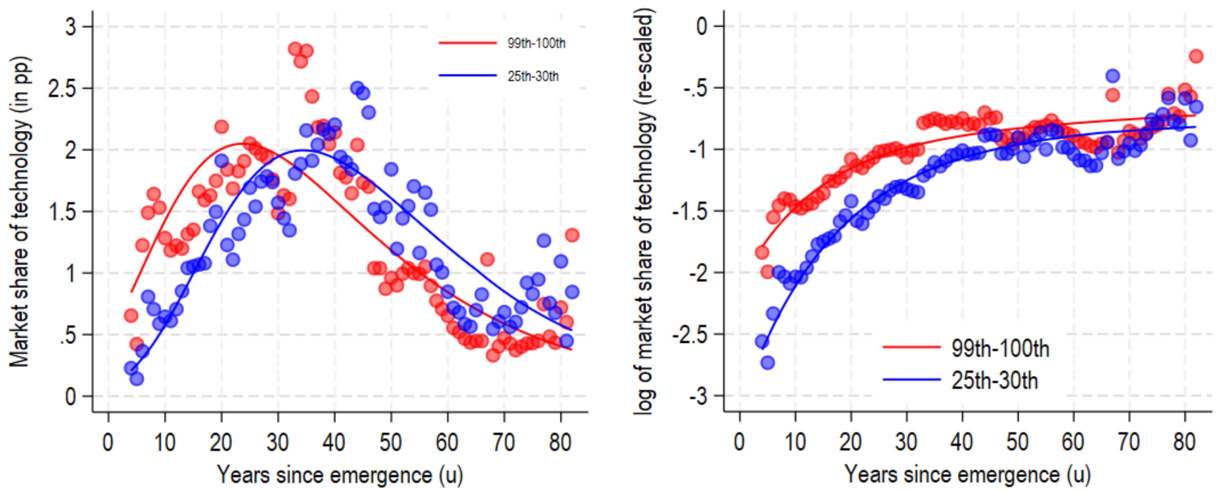

Figure 12: Life cycle of technologies employment shares across locations of different densities. Panel A plots the share of employment per technology of age u (including both college and non-college workers). This is reported as a fraction of employment in technologies under 82 years old and is done separately for low (in blue) and high-density (in red) regions. Panel B plots the left side of the estimating equation (10), $\frac{1}{\sigma-1}\ln \text{Employment share}(u,d) + gu - \frac{1}{\sigma-1}\ln\kappa(\alpha(u), W_h(d), W_\ell(d))$ for both regions. Solid black lines show the fitted values from the non-linear least squares estimation of (10).

adjusting for differences in technology costs due to their skill mix, which depend on the wage rates, $W_h(d)$ and $W_\ell(d)$, faced by firms in each location.

To match the pattern in Panel A of Figure 12, we parametrize $z(u)$ as before and set

$$\ln p(u,d) \;=\; -\phi_0\,(1-d)^{\phi_1}\;\max\{1-u/\bar{u},0\}\;e^{-\phi_2 u}$$

This specification implies that, upon emergence, a technology reaches the highest-density locations ($d=1$) immediately, while it reaches lower-density locations with probability $e^{-\phi_0(1-d)^{\phi_1}}$. Here, ϕ_0 controls the arrival gap between the highest and lowest density locations, and ϕ_1 gives the slope of this arrival rate as a function of density. As technology ages, the probability of arrival increases roughly log-linearly in u, with ϕ_2 accounting for any deceleration in this process. The diffusion process then ends at $\bar{u}$.[19]

[19]This specification can be micro-founded by imagining that every technology comes with an ease-of-adoption term $\nu \geq 0$, distributed exponentially, and a technology of age u is adopted if

$$\nu \geq C(D)\,g(u).$$

Taking logs in (9), rearranging, and bringing in the specification for $z(u)$ and $p(u, d)$ yields the estimating equation

$$\frac{1}{\sigma-1}\ln \text{Employment share}(u,d) + gu - \frac{1}{\sigma-1}\ln \kappa(\alpha(u), \bar{W}_h(d), \bar{W}_\ell(d))$$
$$= \frac{1}{\sigma-1}\ln \text{constant}(d) + g_m u + \frac{1}{\lambda}(g_M - g_m)e^{-\lambda u}$$
$$- \frac{1}{\sigma-1}\,\phi_0\,(1-d)^{\phi_1}\,\max\{1-u/\bar{u}, 0\}\,e^{-\phi_2 u} + \varepsilon_{u,d}, \qquad (10)$$

where we added a measurement error $\varepsilon_{u,d}$ and the location-specific constant in (9). The term adjusting for the skill mix on the left is computed using observed wage rates across density bins from the 2010 American Community Survey and our estimates for $\alpha(u)$.

This equation shows that $z(u)$ and $p(u, d)$ can be recovered from employment shares, adjusting for the growth of frontier technologies and differences in skill mixes. Market shares in high-density places identify $z(u)$. Differences in market shares by age across densities identify $p(u, d)$.[20]

We estimate equation (10) by aggregating data on market shares by technology age across 21 density bins, each containing 5 percent of the commuting zones in the sample, and a bin for the top 0.5% by density. The right-hand side in (8) is shown in Panel B of Figure 5, for both high and lower-density bins. We set $\bar{u} = 65$ on the right-hand side, which provides a sufficiently long window for technologies to fully diffuse, and explore the implications of a longer window in the appendix.

Fitting (10) via non-linear least squares yields $g_m = 0.002$ (se= 0.001), $g_M = 0.081$ (se= 0.018), and $\lambda = 0.073$ (se= 0.012). The estimates differ from those above because they separate the role

The term $C(D) \geq 0$ is an adoption cost as a function of density D. The term $g(u)$ is a decreasing function from $g(0) = 1$ to $g(\bar{u}) = 0$ that captures the decline in adoption costs as the technology ages. The functional form for the first term results from assuming that (i) $C(D)$ is a power function and (ii) density D follows a Pareto distribution (so that $D = D_{min}\,(1-d)^{-1/\alpha}$, where α is the tail index). The functional form for the second term provides a good fit to the data, where the gap (in logs) between the market shares of technologies of age u in dense vs other locations decreases almost log-linearly in u, then tapers off after 60–70 years.

[20] Our specification assumes that technologies reach high density locations immediately. This is an inconsequential normalization. Indeed, if $p(u, 1) < 1$, one can fold $p(u, 1)$ and $z(u)$ together and reinterpret the estimates as relative probabilities $p(u, d)/p(u, 1)$, which is what matters for our quantitative results. Relatedly, our specification assumes all technologies eventually diffuse to all locations. One could assume a fraction $\bar{p}(d)$ of technologies can potentially reach locations with density d. This leads to the exact same estimating equation (with $\bar{p}(d)$ folded into the density-specific constant) and yields identical quantitative results, since the college premium in a location is independent of $\bar{p}(d)$ —a corollary of Proposition 1.

of spatial diffusion from the productivity life cycle $z(u)$. We also obtain $\phi_0 = 2.859$ (se=0.409), $\phi_1 = 0.403$ (se=0.119), and $\phi_2 = 0.035$ (se=0.011). The model fits the data well, with an R^2 of 90.5%. The fitted values are shown by the solid lines in Figure 12. As intended, the model matches the life cycle of employment shares by technology age.

Figure 13 plots the estimated diffusion probabilities, $p(u, d)$. The left panel plots the log of relative arrival probabilities $\ln p(u, d) - \ln p(u, d_{top})$ for each density bin in our sample, relative to the highest bin. The figure reveals large spatial differences in the arrival rate of technology. Once introduced, there is a 200-log-point gap in the probability that a technology reaches the highest versus the lowest density locations (controlled by ϕ_0). Other locations are in between, with ϕ_1 controlling where. These arrival gaps shrink over time until the diffusion process stops. The gaps generate sizable shifts in the market shares of technologies across locations, shown in the right panel.

Panel A: Fitted probability of diffusion $p(u, d)$ *Panel B: Fitted Employment share(u, d)*

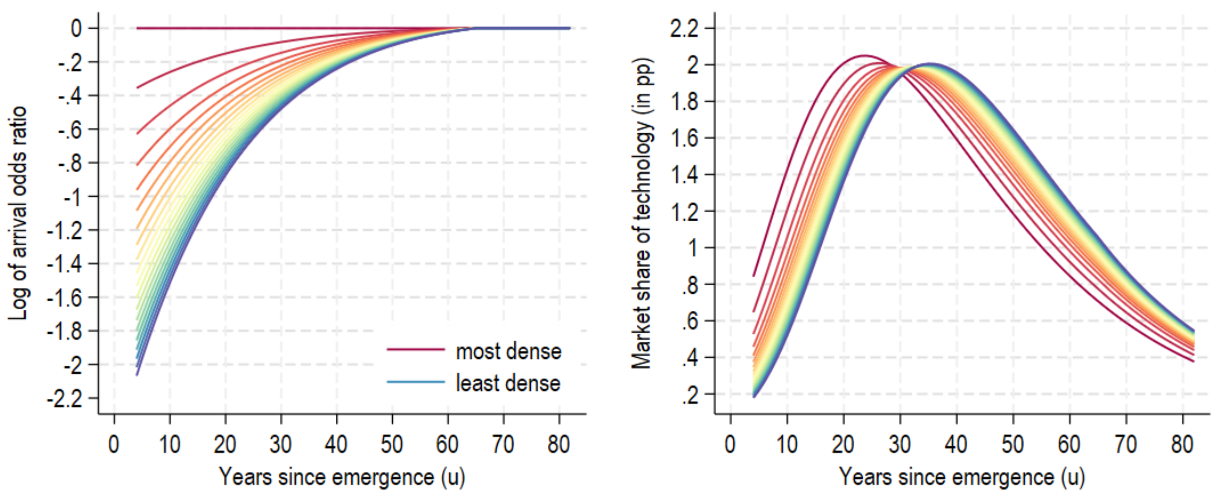

Figure 13: Estimated diffusion probability $p(u, d)$ and implied employment shares. The figure plots estimates from equation (10). Panel A plots fitted probabilities $p(u, d)$ relative to the top density bin in the sample (reported in logs). Panel B plots implied employment shares along a balanced growth path. Red tones indicate higher density bins and blue tones lower density ones.

Table A3 in the Appendix reports a series of robustness checks for our estimates of (10). The table reports estimates with $\bar{u} = 75$, allowing more time for the diffusion process to complete. It also reports estimates excluding ICT technologies and separately for the early and late years in the sample. The table also reports estimates obtained by aggregating the data in density bins, cohort (b), and calendar year (t) cells. As before, we provide estimates that control for year-fixed effects and cohort effects. The estimated diffusion patterns, both as technologies age and across densities, are quite stable across specifications.

4.3 The pace of technology creation and the skill premium across US regions:

As before, we initialize the economy on its balanced growth path and feed in the observed changes in the pace of technology creation, $m(b)$, measured over 1970–2007. All other determinants of the college premium are held constant across time and locations. In particular, we set a common relative supply of skills, $h/\ell = 0.53$, for all locations and years.

Because earlier Census data do not reliably cover low-density locations, we treat the 1980 Census as corresponding to the model's balanced growth path.[21] We calibrate a common limit-obsolescence rate across all locations of 0.2 percent to generate a balanced-growth-path college premium of 44.5 log points for commuting zones in the 95th–99.5th percentiles of the population-density distribution, matching the level in the 1980 Census. All remaining variation in the college premium across regions and time is untargeted.

Figure 14 compares the model-implied college premium (solid lines) to the data (dashed lines). By construction, the model matches the 44.5 log-point college premium in the 95th–99.5th density bin in 1980. Along the balanced growth path, it generates a college premium of 35 log points in low-density locations (32 in the data) and 47.6 log points in the densest locations (47.2 in the data). The estimated differences in $p(u, d)$ account for 13 of the 15 log-point gap in the skill premium between high- and low-density regions observed in 1980, prior to the subsequent rise in inequality.

The increased pace of technology creation leads to a larger increase in the college premium in dense regions. By 2005, the model predicts an increase of 31.1 log points in the highest-density locations and 24.9 log points in the lowest-density locations, compared with 33.5 and 24.8 in the data. The slow spatial diffusion of technologies introduced over 1970–2000 therefore accounts for 6.2 of the 8.7 log-point differential increase in the skill premium between high- and low-density

[21] This is not an issue here, since the effects of the post-1970 acceleration in technology creation unfold gradually and are not noticeable by 1980.

College premium across regions, model vs. data

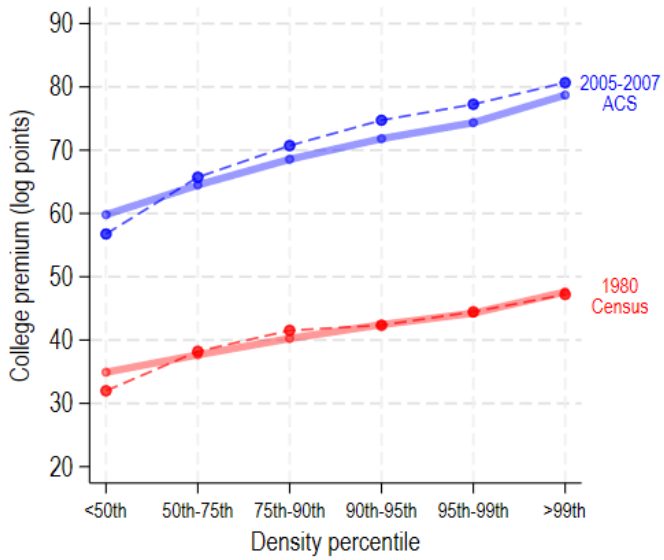

Figure 14: Skill premium across regions: model and data. The figure plots the college premium by population density. The dashed lines plot data for the 1980 Census (in red) and the 2005–2007 American Community Survey (in blue). The solid lines show the model predicted levels for the BGP and 2006.

regions over 1980–2005.

5 The Skill Premium by Worker Age

We next examine how the pace of technology creation affects the skill premium across worker age groups. Figure 15 plots the college premium by people's potential experience, defined as years since completion of schooling. As in Card and Lemieux (2001), the college premium rises first among young workers with 0–10 years of experience. For older workers, with 40–50 years of experience, the increase is muted and protracted.

This section shows that, under the natural assumption that younger workers have a comparative advantage in learning new technologies, our framework also explains this pattern.

5.1 Extending the model to account for demographics

We return to the baseline economy with a single location, but now account for worker demographics. The model describes what happens when workers enter the labor market at age (experience)

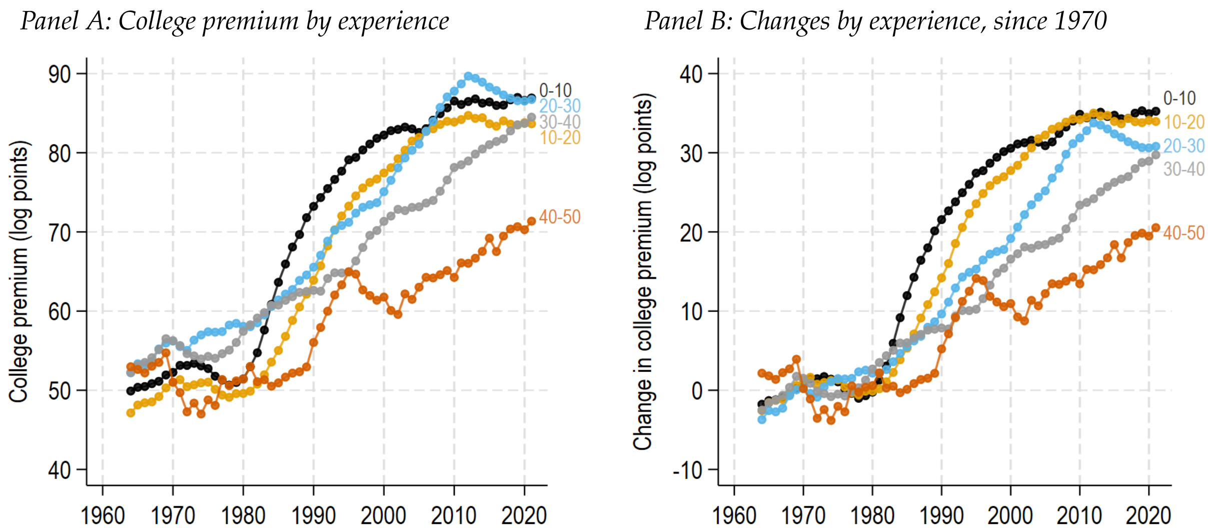

Figure 15: The skill premium by worker age. The figure shows the wage college premium by experience (years since schooling completed). Panel A reports the series in log points. Panel B reports the changes since 1970. Data from the Current Population Survey.

$u_p = 0$ with a fixed level of skill.

Population grows exponentially at rate n. At each point, b, a cohort of people of size

$$N(b) = e^{n\,b}$$

is born. People are born with age $u_p = 0$, age, and retire at age $\bar{u}_p$. A share h of entrants is high-skill and a share ℓ low-skill.

The production side is the same as in the baseline economy. The only difference is that now, the high-skill labor input used per technology of age u is

$$h_t(u) = \int_0^{\bar{u}_p} a(u, u_p)\; h_t(u, u_p)\; e^{\beta_h\, u_p}\; du_p$$

and the low-skill input is

$$\ell_t(u) = \int_0^{\bar{u}_p} a(u, u_p)\; \ell_t(u, u_p)\; e^{\beta_\ell\, u_p}\; du_p.$$

Here $h_t(u, u_p)$ and $\ell_t(u, u_p)$ denote the mass of workers of age u_p employed per technology of age

u. The terms $e^{\beta_h u_p}$ and $e^{\beta_\ell u_p}$ capture the gains from experience, which matter for the life cycle growth of wages, but do not affect specialization patterns.

The function $a(u, u_p)$ gives the productivity of workers of age u_p in technologies of age u. This function is log supermodular in $\langle u, u_p \rangle$, so younger workers have a comparative advantage in new technologies.

To facilitate the characterization of the equilibrium, assume technologies disappear at age $\bar{u}$, which can be chosen arbitrarily large. Labor-market clearing requires that workers in each age group $u_p \in [0, \bar{u}_p]$ are allocated across vintages $u \in [0, \bar{u}]$

$$\int_0^{\bar{u}} m_t(u)\, h_t(u, u_p)\, du \;=\; h\, e^{n\,(t-u_p)}, \quad \int_0^{\bar{u}} m_t(u)\, \ell_t(u, u_p)\, du \;=\; \ell\, e^{n\,(t-u_p)}.$$

Equilibrium and effects of changes in the pace of technology creation: As in Costinot and Vogel (2010), the equilibrium features specialization by age: young workers specialize in new technologies, older workers in older ones. This is described by strictly increasing matching functions $\mathcal{U}_{h,t}(u_p)$ and $\mathcal{U}_{\ell,t}(u_p)$ that give the technology age used by workers of age u_p. These can change over time, depending on the pace of technology creation.

Given the path of technology vintages $\{m_t(u)\}$, an equilibrium is given by a sequence of assignment functions $\mathcal{U}_{\ell,t}(u_p)$, a sequence for output Y_t and real wages $\{W_{h,t}(u_p), W_{\ell,t}(u_p)\}$, that vary over time—due to changes in the pace of technology creation—and workers' age—due to differences in comparative advantage.

The equilibrium conditions follow Costinot and Vogel (2010), and are provided in the appendix. As in our baseline model, our first proposition shows that the model admits a BGP with exponential growth and constant skill premia by age.

Proposition 5 (Balanced Growth Path and Demographics)**.** *Suppose the pace of technology creation is constant, i.e. $m(b) = m$, so that $m_t(u) = m$. There exists a balanced growth path along which real wages and output per worker grow at a rate g, young workers specialize in new technologies, and the skill premium by worker age and the assignment functions are constant in time and independent of m.*

The proposition shows that the model admits a balanced growth path in which younger workers specialize in using new technologies, while older workers use older technologies.

Proposition 6 (Changes in the Pace of Technology Creation and Effects across Age Groups)**.** *Assume $\sigma = \gamma$ and $\alpha(u)$ is sufficiently high for recently introduced technologies. Consider an economy in its balanced growth path. An increase in $m(b)$ from m to $m' > m$, whether permanent or temporary, generates a transitory increase in the college premium for all worker age groups. The increase is more front-loaded among young workers: for any $u_p < u_p'$, the increase in the skill premium at age u_p exceeds the increase in the college premium at age u_p' early on in the transition.*

The proposition shows that an increase in the pace of technology creation raises the skill premium, with a stronger initial effect on the young. This is because the large cohort of technologies introduced at t_0 demand skilled young workers, bidding up their wages. The skill premium rises later for older workers, who are slower to adopt new technologies.

The requirement that $\sigma = \gamma$ is imposed for tractability, as it simplifies the analysis. The requirement that $\alpha(u)$ is sufficiently high for new technologies ensures that young, skilled workers are more exposed to new technologies than other young workers.

Propositions 5 and 6 clarify the link between the pace of technology creation and changes in the skill premium across worker age groups. We turn to the data to examine these implications and quantify the contribution of this mechanism to the trends in Figure 15.

5.2 Estimating the productivity schedule $a(u, u_p)$

For this exercise, we build on the baseline calibration in Section 2 and introduce estimates of the returns to experience, β_h and β_ℓ, as well as the function $a(u, u_p)$. Using CPS data over 1975–2021, we estimate yearly returns to experience of $\beta_h = 1.16\%$ for college-educated workers and $\beta_\ell = 1.01\%$ for non-college workers using Mincer regressions.

The function $a(u, u_p)$ governs younger workers' comparative advantage in new technologies and is central to generating differential effects of technological innovation on the skill premium across age groups. We calibrate $a(u, u_p)$ using data on computer use at work by worker age from the Current Population Survey (CPS). This dataset is well-suited for this purpose because it provides consistent information on technology use across age groups. We assume that age patterns in computer use generalize to other workplace technologies.

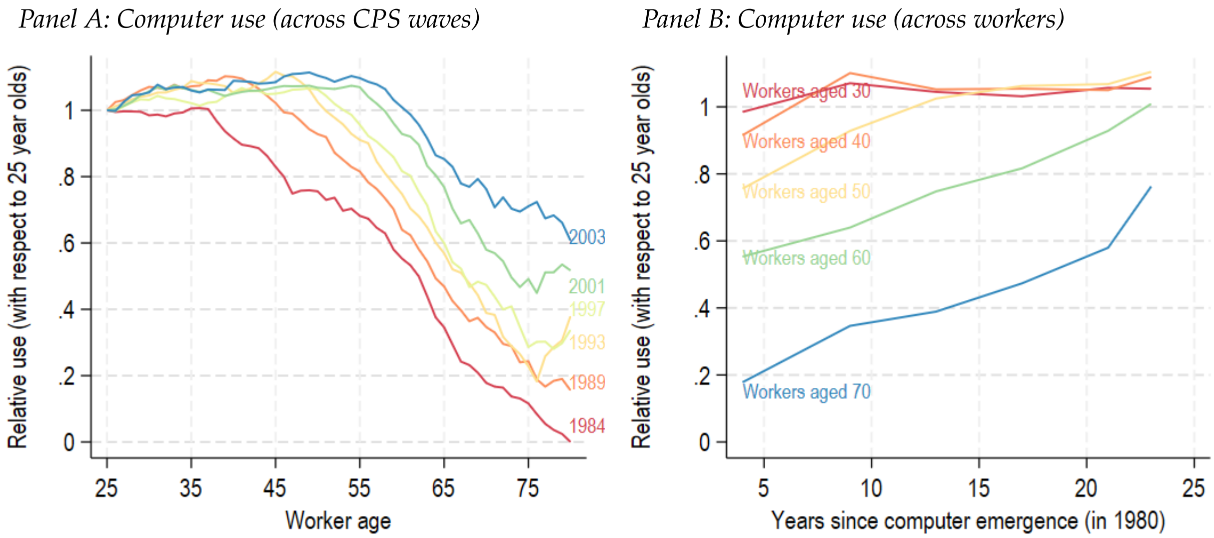

Figure 16: Computer use at work by age. The figure plots data on computer use from the Current Population Survey (CPS). Panel A reports computer use by worker age (on the horizontal axis) across CPS waves. Panel B reports computer use by technology age (on the horizontal axis, assuming computers emerged in 1981) by worker age. Both panels report use relative to 25-year-olds.

Panel A of Figure 16 plots computer use at work by worker age across CPS waves. To aid interpretation, we normalize usage in each age group to the share of 25-year-olds who use computers. In early waves, computer use is heavily concentrated among younger workers. Over time, usage converges, with the largest increases occurring among older workers. Panel B presents the same pattern, now as a function of technology age. The horizontal axis reports years since workplace computers were introduced in 1981, and the vertical axis plots relative usage for 70-, 60-, 50-, 40-, and 30-year-olds, normalized by usage among 25-year-olds.[22]

We assume that reported computer-use rates in the CPS proxy for workers' proficiency in using the technology. In particular, we posit

$$\text{share users}(u, u_p) = M\, a(u, u_p)\, e^{\varepsilon_{u,u_p}},$$

[22]Various workplace computers were introduced around 1981, including the Apple II and the IBM Personal Computer. Our approach dates the emergence of personal computers to 1985, four years later, reflecting the timing at which patenting activity associated with personal computers becomes sufficiently widespread to meet our definition of emergence (see Table 1).

where $M > 0$ is a normalization constant and ε_{u,u_p} is an error term orthogonal to worker age and technology age. We parameterize

$$\ln a(u, u_p) = -\rho_0 \, u_p^{\rho_1} \, e^{-\rho_2 u},$$

and map workers of age 25 to $u_p = 0$ in the model. We also set $\bar{u}_p = 50$, so that workers retire after 50 years. This specification assumes that 25-year-old workers have unit productivity at new technologies and older workers face an initial productivity gap of $\rho_0 \, u_p^{\rho_1}$, which shrinks as the technology ages at rate ρ_2. This specification is designed to fit the patterns in Panel B of Figure 16.

Computer use by worker age, data and fitted curves

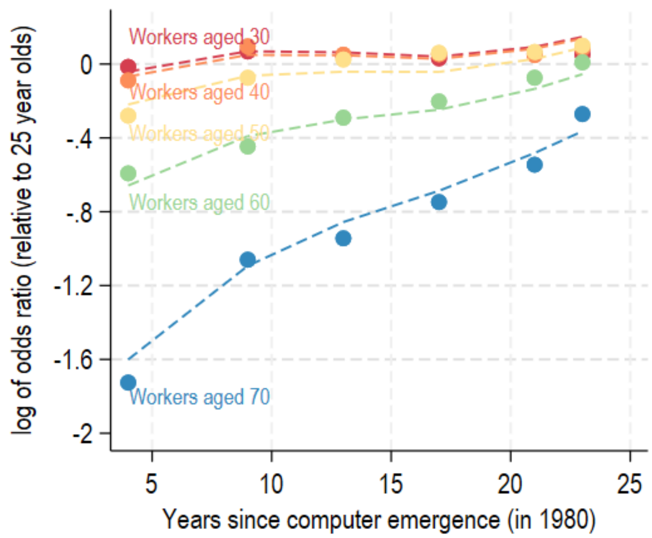

Figure 17: Fitted model for $a(u, u_p)$ based on computer use data. The figure reports the model fit for equation (11). The dots are data from the Current Population Survey on computer use (relative to 25-year-olds, and in logs, as on the left-hand side of (11)). The solid lines plot the fitted curves.

These assumptions allow us to estimate $a(u, u_p)$ from the model

$$\ln \frac{\text{share users}(u, u_p)}{\text{share 25-year old users}(u)} = -\rho_0 \, u_p^{\rho_1} \, e^{-\rho_2 u} + \underbrace{\varepsilon_{u,u_p} - \varepsilon_{u,0}}_{v_{u,u_p}}. \tag{11}$$

Estimating this equation via NLLS, controlling also for CPS wave effects, yields $\rho_0 = 2.9$ (se=0.1), $\rho_1 = 3.7$ (se=0.1), and a low catching-up rate of $\rho_2 = 0.059$ (se=0.003). Figure 17 plots the left-hand side of this equation and the fitted curves. As intended, our parametrization fits the data well, yielding an R^2 of 97.6%.

5.3 The pace of technology creation and the skill premium across age groups:

As in our previous exercises, we initialize the economy on its balanced growth path and feed in observed changes in the pace of technology creation, $m(b)$, over 1970–2007.

Figure 18 reports the changes in the college premium generated by the model for workers of different ages, both in levels (Panel A) and relative to 1970 (Panel B). Because college-educated workers have a higher return to their experience than non-college ones, the baseline level of the college premium is similar across age groups in the economy's BGP, consistent with the 1970 data. In response to the acceleration in the pace of technology creation over 1970–2000, the model reproduces the age-specific patterns observed in the data, with the college premium rising first for younger workers and later—and much more modestly—for older workers.

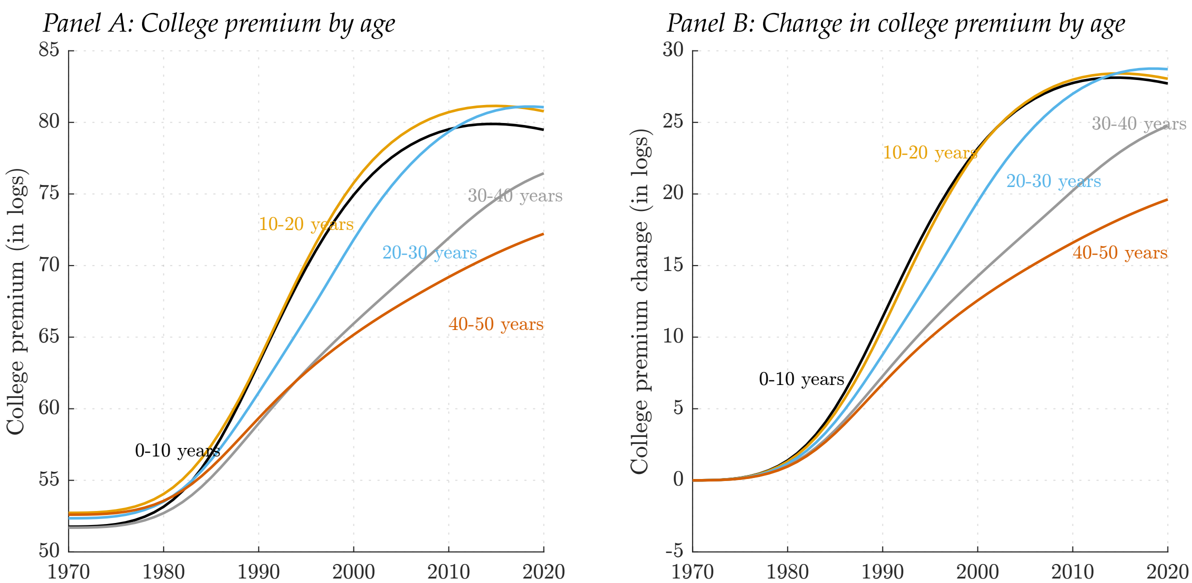

Figure 18: Model estimates of the effects of the pace of technology on the skill premium by worker experience. The figure shows the model results for the college premium by worker experience. Panel A reports the series in log points. Panel B reports changes since 1970.

The college premium for workers with 0-10 years of experience rises by 28 log points by 2020, while the college premium for those with 40–50 years of experience rises by 20 log points. The 8 point gap accounts for 53% of the gap in the data, where the college premium rises by 35 log points for young workers and 20 log points for older ones.

6 Concluding Remarks

This paper argues that changes in the pace of technology creation are a key driver of the rise in the skill premium over time, across regions, and age groups. An acceleration in the pace of technology creation affects the skill premium along these dimensions because: (i) college-educated workers have an advantage in using new technologies (Schultz, 1975); (ii) new technologies arrive earlier in dense cities; and (iii) young workers are early adopters.

A key contribution of the paper is to leverage advances in text-to-data methods (building on Kalyani et al., 2025) to quantify these forces. Measured changes in the pace of technology creation generate a 28 log-point increase (32 percent) in the college premium and account for one-third of the growth in demand for skills since 1970. Across regions, the mechanism accounts for 70% of the urban bias of the skill premium increase. Across age groups, a calibrated version of the model that matches computer-use patterns by worker age accounts for 53% of the age-specific trends in the college premium in the data.

References

Acemoglu, D. and D. H. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter and D. Card (Eds.), Handbook of Labor Economics, Volume 4B, pp. 1043–1171. Elsevier.

Acemoglu, D. and P. Restrepo (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. American Economic Review 108(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2022). Tasks, automation, and the rise in u.s. wage inequality. Econometrica 90(5), 1973–2016.

Autor, D. H. (2019, May). Work of the past, work of the future. AEA Papers and Proceedings 109, 1–32.

Autor, D. H., C. Chin, A. Salomons, and B. Seegmiller (2023). New frontiers: The origins and content of new work, 1940–2018. Technical report, Unpublished Manuscript. Working Paper or Forthcoming.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008). Trends in u.s. wage inequality: Revising the revisionists. The Review of Economics and Statistics 90(2), 300–323.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998). Computing inequality: Have computers changed the labor market? The Quarterly Journal of Economics 113(4), 1169–1213.

Bartel, A. P. and F. R. Lichtenberg (1987). The comparative advantage of educated workers in implementing new technology. The Review of Economics and statistics, 1–11.

Beaudry, P., D. A. Green, and B. M. Sand (2016). The great reversal in the demand for skill and cognitive tasks. Journal of Labor Economics 34(S1), S199–S247.

Broda, C. and D. E. Weinstein (2006, 05). Globalization and the gains from variety. The Quarterly Journal of Economics 121(2), 541–585.

Burke, S. N. and C. A. Barnes (2006). Neural plasticity in the ageing brain. Nature Reviews Neuroscience 7(1), 30–40.

Burstein, A., E. Morales, and J. Vogel (2019). Changes in between-group inequality: Computers, occupations, and international trade. American Economic Journal: Macroeconomics 11(2), 289–326.

Card, D. and T. Lemieux (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis. The Quarterly Journal of Economics 116(2), 705–746.

Caselli, F. (1999, March). Technological revolutions. American Economic Review 89(1), 78–102.

Costinot, A. and J. Vogel (2010). Matching and inequality in the world economy. Journal of Political Economy 118(4), 747–786.

Davies, M. (2010). The corpus of historical american english (coha). Online database, https://www.english-corpora.org/coha/. Accessed YYYY-MM-DD.

De Loecker, J. (2011). Product differentiation, multi-product firms and estimating the impact of trade liberalization on productivity. Econometrica 79(5), 1407–1451.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2016). Where has all the skewness gone? the decline in high-growth (young) firms in the us. European Economic Review 86, 4–23.

Doms, M., T. Dunne, and K. R. Troske (1997, 02). Workers, wages, and technology. The Quarterly Journal of Economics 112(1), 253–290.

Eckert, F., S. Ganapati, and C. Walsh (2022). Urban-biased growth: a macroeconomic analysis. Technical report, National Bureau of Economic Research.

Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco.

Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? American Economic Review 98(1), 394–425.

Galor, O. and O. Moav (2000, 05). Ability-biased technological transition, wage inequality, and economic growth. The Quarterly Journal of Economics 115(2), 469–497.

Glaeser, E. L. (1999). Learning in cities. Journal of Urban Economics 46(2), 254–277.

Greenwood, J. and M. Yorukoglu (1997). 1974. Carnegie-Rochester Conference Series on Public Policy 46, 49–95.

Horn, J. L. and R. B. Cattell (1967). Age differences in fluid and crystallized intelligence. Acta Psychologica 26, 107–129.

Kalyani, A. (2024). The Creativity Decline: Evidence from US Patents. Federal Reserve Bank of St. Louis, Research Division.

Kalyani, A., N. Bloom, M. Carvalho, T. A. Hassan, J. Lerner, and A. Tahoun (2025). The diffusion of new technologies. Technical report, National Bureau of Economic Research.

Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963–1987: Supply and demand factors. The Quarterly Journal of Economics 107(1), 35–78.

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2021, September). Measuring technological innovation over the long run. American Economic Review: Insights 3(3), 303–320.

Kogan, L., L. Schmidt, and B. Seegmiller (2023). Technology and labor displacement: Evidence from linking patents with worker-level data. Technical report, Unpublished Manuscript. Working Paper or Forthcoming.

Kortum, S. and J. Lerner (1999). What is behind the recent surge in patenting? Research policy 28(1), 1–22.

Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. Econometrica 68(5), 1029–1053.

Lehr, N. H. (2023). Innovation in an Aging Economy.

Lerner, J. and A. Seru (2022). The use and misuse of patent data: Issues for finance and beyond. The Review of Financial Studies 35(6), 2667–2704.

Lindenberger, U. and P. B. Baltes (1997). Intellectual functioning in old and very old age: Cross-sectional results from the berlin aging study. Psychology and Aging 12(3), 410–432.

Mukoyama, T. (2004). Diffusion and innovation of new technologies under skill heterogeneity. Journal of Economic Growth 9(4), 451–479.

Oberfield, E. and D. Raval (2021). Micro data and macro technology. Econometrica 89(2), 703–732.

Rubinton, H. (2020). The geography of business dynamism and skill biased technical change. FRB St. Louis Working Paper (2020-20).

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. Psychological Review 103(3), 403–428.

Schultz, T. W. (1975, September). The Value of the Ability to Deal with Disequilibria. Journal of Economic Literature 13(3), 827–846.

Syverson, C. (2004). Market structure and productivity: A concrete example. Journal of Political Economy 112(6), 1181–1222.

Appendix for "The Skill Premium in Times of Rapid Technological Change"

By Tarek Hassan, Aakash Kalyani, and Pascual Restrepo.

A Data construction details

Current Population Survey data: we use CPS data for Figures 1, and 15.

We use data from the March Current Population Survey (CPS) for the period 1965–2021. The raw CPS files were obtained from the NBER, and we applied standard sample restrictions and cleaning procedures.

We compute weekly wages by dividing total labor income by the number of weeks worked during the year. We convert these into real weekly wages using the Bureau of Economic Analysis personal consumption expenditure (PCE) deflator. To address top coding, we replace top-coded income observations with 1.5 times the reported top-coded value.

We define skilled workers as those with a college or post-college degree. Workers with some college education but no completed degree are classified as low-skill.

To account for changes over time in the composition of college-educated and non-college workers, we compute weekly wages by detailed demographic group. Groups are defined by gender, age (five age categories), race, and detailed educational attainment (no high school, high school degree, some college, college degree, and post-college degree). We then construct compositionally adjusted wages by averaging group-level wages weighted by each group's 1980 population share. This ensures that our series for the college premium reflects wage changes for a fixed, representative composition of college and non-college workers.

We measure labor supply in a manner consistent with total reported labor income. Specifically, we compute the compositionally adjusted supply of college labor (measured in weeks worked) as total labor income earned by college-educated workers divided by their compositionally adjusted weekly wage. We apply the same procedure to compute the supply of non-college labor. Figure A1 shows our series for the relative supply of college-educated labor, $\frac{h_t}{\ell_t}$.

In addition to the main CPS files, we use the Computer and Internet Use Supplement, which

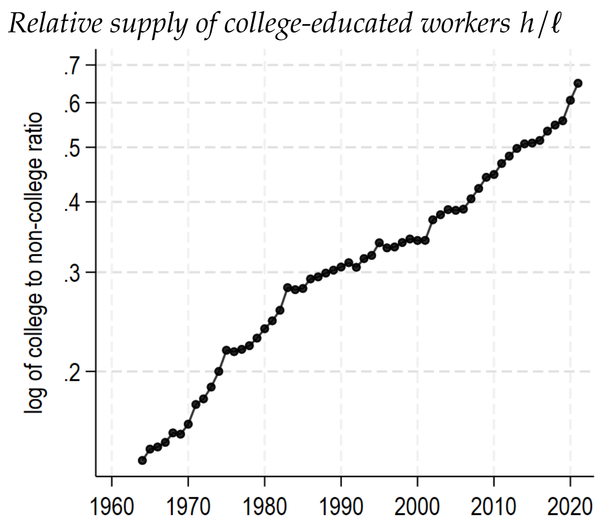

Figure A1: Relative supply of college-educated workers in the US, 1960–2022. The figure plots the ratio of hours worked by college-educated workers relative to non-college-educated workers in the United States from 1965 to 2021, computed using data from the Current Population Survey (CPS). College-educated workers are defined as those with a bachelor's degree or higher. The vertical axis is on a logarithmic scale.

provides information on computer use at work. The relevant question on computer use is available in the 1984, 1989, 1993, 1997, 2001, and 2003 waves.

For additional details on data construction and sample definitions, we refer the reader to the replication kit accompanying the paper.

Census and American Community Survey data: we use the 1980 Census and 2005–2007 ACS data for Figure 10.

We use data from the 1980 Decennial Census and the 2005–2007 multiyear American Community Survey (ACS). The raw microdata were obtained from IPUMS USA, and we applied standard sample restrictions and cleaning procedures.

We compute weekly wages by dividing total labor income by the number of weeks worked during the year, and convert these to real weekly wages using the Bureau of Economic Analysis personal consumption expenditure (PCE) deflator. As in the CPS, we address top coding by replacing top-coded income observations with 1.5 times the reported value.

We define skilled workers as those with a college or post-college degree, and classify workers with some college education but no completed degree as low-skill.

To account for changes in the composition of college-educated and non-college workers, we compute weekly wages by detailed demographic group, defined by gender, age (five age categories), race, and detailed educational attainment (no high school, high school degree, some college, college degree, and post-college degree). We then construct compositionally adjusted wages by averaging group-level wages weighted by each group's 1980 population share. This procedure ensures that our measures of the college premium capture wage differences for a fixed, representative composition of workers.

We compute these compositionally adjusted wages separately for each of 722 commuting zones. We link observations across the Census and ACS using commuting-zone crosswalks constructed by David Dorn.

For additional details on data construction and geographic crosswalks, we refer the reader to the replication kit accompanying the paper.

Patent text. We download full text of all patents filed by inventors who report a US location in the USPTO from 1930-2014. These text are processed in two parts: 1) patents filed between 1976 and 2014 are all available in digitized bulk downloads from the USPTO, 2) patents filed between 1930 and 1975 are available in pdf image formats and are processed using optical character recognition and converted to digitized text

B Robustness checks for the estimates of $\alpha(u)$, $z(u)$, and $p(d,u)$.

Estimates of $\alpha(u)$: Table A1 presents our estimates for $\alpha(u)$ and robustness checks. Column 1 reproduces our baseline estimates of equation (6). Column 2 removes ICT technologies from the sample before aggregating the data. Columns 3 and 4 re-estimate equation (6) using only the job-posting data for early (2010–2015) and late years (2016–2023) in our job-posting sample.

The remaining columns present estimates of $\alpha(u)$ from a panel variant of equation (6):

$$\ln\left(\frac{h_t(u)}{l_t(u)}\right) + \gamma \ln\left(\frac{W_{t,h}}{W_{t,\ell}}\right) = \theta_0 - \theta_1 u + \epsilon_{u,t}. \tag{A1}$$

Table A1: **Estimates of $\alpha(u)$ and robustness checks.**

	Baseline estimates, equation (6)	Removing ICT technologies	Job-posting data for 2010 – 2015	Job-posting data for 2016 – 2023	Estimates of (A1) at the cohort × year: Panel estimates	Time fixed effect	Cohort effects (20-year cohorts)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age coefficient, -θ_1	–0.012***	–0.011***	–0.010***	–0.012***	–0.013***	–0.012***	–0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Constant term θ_0	1.260***	1.193***	1.280***	1.227***	1.315***	1.264***	1.328***
	(0.020)	(0.024)	(0.029)	(0.026)	(0.022)	(0.021)	(0.024)
R-squared	0.908	0.862	0.790	0.898	0.610	0.641	0.654
N	79	79	71	73	924	924	924

Notes: The Table reports OLS estimates of equation (6) (columns 1-4, estimated by aggregating technologies by age u) and (A1) (columns 5-7, estimated by aggregating technologies by cohort b and year t). These equations describe how the demand for skills declines as a technology ages. In columns 5–7, each cohort–year cell is weighted by the inverse of the number of cells for technologies of age u. Robust standard errors are reported in parentheses.

In this equation, $h_t(u)$ and $l_t(u)$ denote the total number of college and non-college job postings, respectively, associated with technologies of age u in year t, with t given by the years in the Lighcast data, $t = 2010, \dots, 2023$. The difference with respect to the estimating equation in the main text is that here we aggregate the job-posting data to a panel of cohorts × year cells.

Column 5 reports estimates of equation (A1). In this specification, each cell is weighted by the inverse of the number of same-age technology cells, ensuring that each cohort contributes an equal weight. Column 6 reports estimates that control for time-fixed effects. Finally, column 7 presents estimates controlling for 20-year cohort dummies, 1940–1960, 1960–1980, 1980–2000, and 2000-2020.

Estimates of $z(u)$: Table A2 presents our estimates for $z(u)$ and robustness checks. Column 1 reproduces our baseline estimates of equation (8). Column 2 removes ICT technologies from the sample before aggregating the data. Columns 3 and 4 re-estimate equation (8) using only the job-posting data for early (2010–2015) and late years (2016–2023) in our job-posting sample.

Table A2: **Estimates of $z(u)$ and robustness checks.**

					Estimates of (A2) at the cohort × year		
	Baseline estimates, equation (8)	Removing ICT technologies	Job-posting data for 2010 – 2015	Job-posting data for 2016 – 2023	Panel estimates	Time fixed effect	Cohort effects (20-year cohorts)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
g_m	0.002	–0.018	–0.007	0.001	–0.006	–0.015	0.003
	(0.003)	(0.025)	(0.009)	(0.004)	(0.010)	(0.016)	(0.008)
g_M	0.082***	0.057***	0.080***	0.119***	0.089***	0.074***	0.069*
	(0.010)	(0.007)	(0.015)	(0.035)	(0.025)	(0.019)	(0.036)
λ	0.048***	0.016	0.041**	0.055***	0.046**	0.033*	0.066
	(0.009)	(0.010)	(0.016)	(0.015)	(0.021)	(0.018)	(0.045)
R-squared	0.972	0.976	0.853	0.924	0.405	0.423	0.420
N	79	79	71	73	924	924	924

Notes: The Table reports NLLS estimates of equation (8) (columns 1-4, estimated by aggregating technologies by age u) and (A2) (columns 5-7, estimated by aggregating technologies by cohort b and year t). These equations describe how the market share (in employment) of technology changes as they age. In columns 5-7, each cohort–year cell is weighted by the inverse of the number of cells for technologies of age u. Robust standard errors are reported in parentheses.

The remaining columns present estimates of $z(u)$ from a panel variant of equation (8):

$$\frac{1}{\sigma-1}\ln \text{Employment share}_t(u) + gu - \frac{1}{\sigma-1}\ln \kappa(\alpha(u), W_{h,t}, W_{\ell,t}) = \frac{1}{\sigma-1}\ln \text{constant} + g_m u + \frac{1}{\lambda}(g_M - g_m)e^{-\lambda u} + \varepsilon_{u,t}, \tag{A2}$$

In this equation, Employment share$_t(u)$ denotes the share of employment per technology of age u in year t, with t given by the years in the Lighcast data, $t = 2010, \ldots, 2023$. The difference with respect to the estimating equation in the main text is that here we aggregate the job-posting data to a panel of cohorts × year cells.

Column 5 reports estimates of equation (A2). In this specification, each cell is weighted by the inverse of the number of same-age technology cells, ensuring that each cohort contributes an equal weight. Column 6 reports estimates that control for time-fixed effects. Finally, column 7 presents estimates controlling for 20-year cohort dummies, 1940–1960, 1960–1980, 1980–2000, and 2000-2020.

To facilitate comparison across specifications, Figure A2 plots the employment shares by technology age implied by the estimates along the economy's BGP. The figure shows that all specifications produce a similar inverse-U-shaped pattern, with minor differences in the timing of peak employment and in the rate at which employment rises and falls.

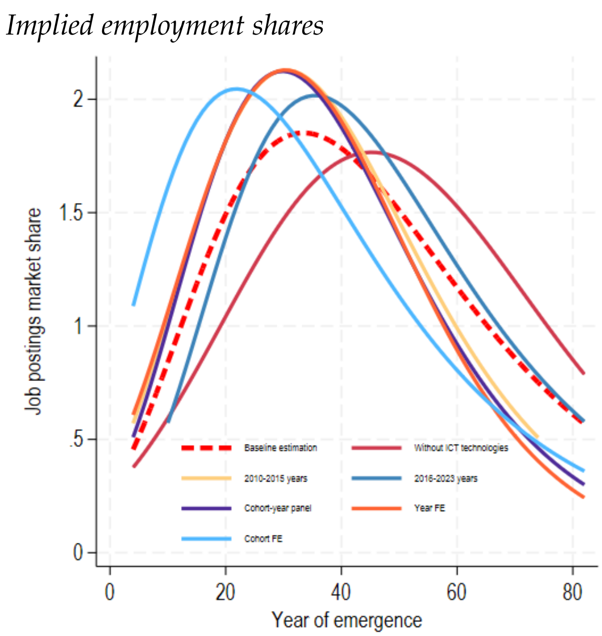

Figure A2: Robustness of $z(u)$ estimates and the implied lifecycle of employment shares. The figure plots the model-implied employment per technology of age u. The curves correspond to the estimates for $z(u)$ in Table A2.

Joint estimates of $z(u)$ and $p(d,u)$: Table A3 presents our joint estimates for $z(u)$ and $p(u,d)$ plus robustness checks. Column 1 reproduces our baseline estimates of equation (10). This specification is estimated at the density bin (using the 21 bins described in the text) and technology age cell. Each bin is weighted by the average annual number of job postings in that density bin. Column 2 presents estimates that set $\bar{u} = 75$, allowing for a longer diffusion window. Column 3 removes ICT technologies from the sample before aggregating the data. Columns 4 and 5 re-estimate equation (10) using only the job-posting data for early (2010–2015) and late years (2016–2023) in our job-posting sample.

Table A3: **Joint estimates of $z(u)$ and $p(u, d)$ plus robustness checks.**

	Baseline estimates, equation (10)	Estimates setting $\bar{u} = 75$ in equation (10)	Removing ICT technologies	Job-posting data for 2010 – 2015	Job-posting data for 2016 – 2023	Estimates of (A3) at the density bin × cohort × year: Panel estimates	Time fixed effect	Cohort effects (20-year cohorts)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimates for $z(u)$:								
g_m	0.002*	0.002	0.006***	–0.002	0.002	–0.002	–0.004	0.006**
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.003)
g_M	0.081***	0.087***	0.063***	0.096***	0.184***	0.089***	0.082***	0.105**
	(0.018)	(0.015)	(0.011)	(0.020)	(0.056)	(0.023)	(0.019)	(0.045)
λ	0.073***	0.073***	0.044***	0.078***	0.093***	0.069***	0.061***	0.136***
	(0.012)	(0.010)	(0.010)	(0.014)	(0.016)	(0.015)	(0.013)	(0.047)
Estimates for $p(d, u)$:								
ϕ_0	2.859***	2.695***	1.920***	2.658***	3.257***	2.754***	2.653***	2.549***
	(0.409)	(0.320)	(0.308)	(0.464)	(0.908)	(0.538)	(0.488)	(0.444)
ϕ_1	0.403***	0.422***	0.334**	0.332**	0.331**	0.437**	0.434**	0.548***
	(0.119)	(0.117)	(0.130)	(0.137)	(0.143)	(0.195)	(0.190)	(0.209)
ϕ_2	0.035***	0.026***	0.032***	0.029**	0.040**	0.035**	0.030*	0.028*
	(0.011)	(0.009)	(0.012)	(0.013)	(0.016)	(0.018)	(0.015)	(0.016)
R-squared	0.905	0.906	0.965	0.794	0.868	0.338	0.352	0.355
N	1,659	1,659	1,659	1,490	1,533	19,106	19,106	19,106

Notes: The Table reports NLLS estimates of equation (10) (columns 1-5, estimated by aggregating technologies at the density bin and age u cell) and (A3) (columns 6-8, estimated by aggregating technologies by density bin, cohort b and year t). These equations describe how the market share (in employment) of technology changes as it ages across locations with different population densities. In columns 1–5, each density bin is weighted by the average yearly job postings in that bin. In columns 6-8, each density bin-cohort–year cell is weighted by the product of the average yearly job postings in that bin and the inverse of the number of cells for technologies of the same age u. Robust standard errors are reported in parentheses.

The remaining columns present estimates of $z(u)$ and $p(u, d)$ from a variant of equation (10):

$$\begin{aligned}\frac{1}{\sigma-1}\ln \text{Employment share}_t(u,d) + gu - \frac{1}{\sigma-1}\ln \kappa(\alpha(u), W_{h,t}(d), W_{\ell,t}(d))\\ = \ln \frac{1}{\sigma-1}\text{constant}(d) + g_m u + \frac{1}{\lambda}(g_M - g_m)e^{-\lambda u}\\ - \frac{1}{\sigma-1}\,\phi_0\,(1-d)^{\phi_1}\,\max\{1-u/\bar{u}, 0\}\,e^{-\phi_2 u} + \varepsilon_{u,t,d},\end{aligned} \tag{A3}$$

In this equation, Employment share$_t(u, d)$ denotes the employment share per technology of age u in density bin d and year t, with t given by the years in the Lighcast data, $t = 2010, \ldots, 2023$. The difference with respect to the estimating equation in the main text is that here we aggregate the data to a panel of density bins × cohorts × year cells.

Column 6 reports estimates of equation (A3). Each cell is weighted by the product of the average yearly job postings in that bin and the inverse of the number of cells for technologies of age u (to give them equal contributions). Column 7 reports estimates that control for time-fixed effects. Finally, column 8 presents estimates controlling for 20-year cohort dummies, 1940–1960, 1960–1980, 1980–2000, and 2000-2020.

To facilitate comparison across specifications, Figure A3 plots the employment shares by technology age implied by the estimates along the economy's BGP for the highest density bin. The figure shows that all specifications produce a similar inverse-U-shaped pattern, with minor differences in the timing of peak employment and in the rate at which employment rises and falls. The right panel shows that the estimated gap in technology arrival across locations, given by $\phi_0\,(1-d)^{\phi_1}$, is also comparable across specifications.

Panel A: Fitted employment shares *Panel B: Fitted arrival gap, $\phi_0\,(1-d)^{\phi_1}$*

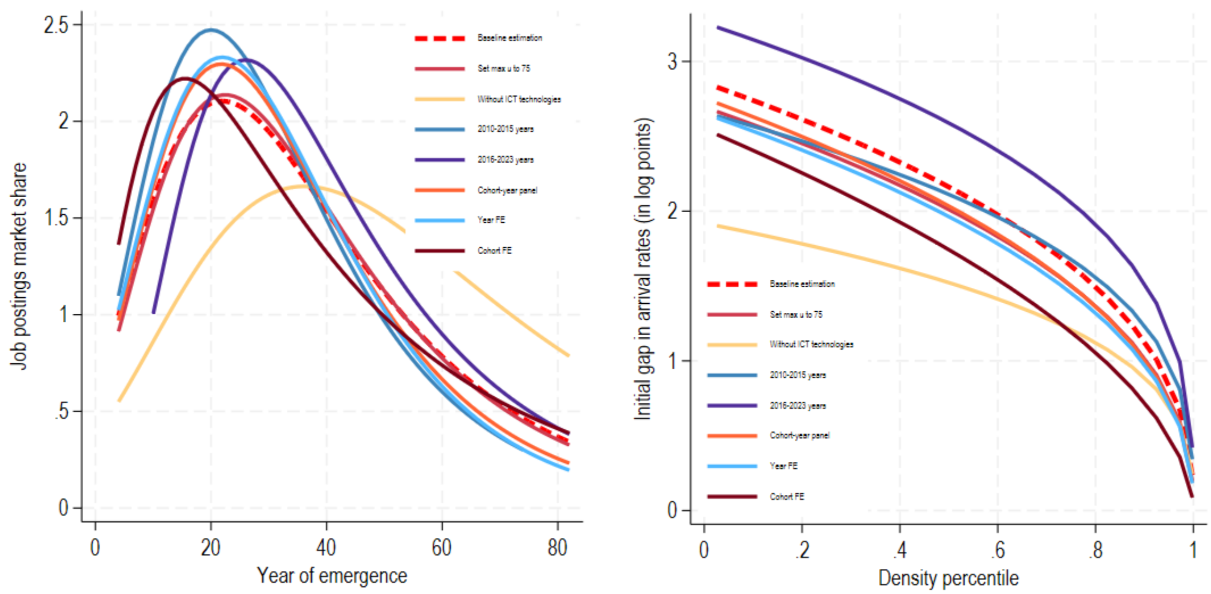

Figure A3: Robustness of $z(u)$ and $p(u,d)$ estimates. The left panel plots the model-implied employment per technology of age u. The right panel plots the initial technology arrival gap (in logs) relative to the highest-density locations. This is given by $\phi_0\,(1-d)^{\phi_1}$. The different curves are obtained using the estimates for $z(u)$ and $p(u,d)$ in Table A3.

Estimates of $\alpha(u)$ by regional density. Table A4 estimates equation (A1) on a panel of density $\times$ cohort $\times$ year cells.

Table A4: **Estimates of $\alpha(u)$ by regional density**

	Pooled OLS (1)	Density FE (2)	High density (3)	Lower density (4)	Heterogeneous slope (5)
Age coefficient, -θ_1	–0.013***	–0.013***	–0.013***	–0.013***	–0.013***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
Age coefficient, -θ_1 · (Density < 95 pct.)					0.000
					(0.001)
Constant term θ_0 · (Density < 95 pct.)					–0.032
					(0.030)
Constant term θ_0	1.057***	1.123***	1.111***	1.054***	1.086***
	(0.011)	(0.008)	(0.021)	(0.012)	(0.027)
R-squared	0.249	0.553	0.261	0.453	0.323
N	6,438	5,541	3,668	3,694	6,438

Notes: The table reports weighted OLS estimates of equation (A1) at the density × cohort × year level. Column (1) reports the baseline specification. Column (2) includes density fixed effects. Columns (3) and (4) estimate the age profile separately in high- and lower-density regions (using the density cut defined in the text). Column (5) allows both the intercept and the age gradient to differ by density via an indicator for (Density < 95 pct.) and its interaction with age. Each cohort–year cell is weighted by the inverse of the number of cohort–year cells for technologies of a given age. Robust standard errors are reported in parentheses.

Column (1) reports the pooled specification, while column (2) adds density fixed effects. Columns (3) and (4) estimate the age profile separately in high-density and lower-density regions (using the density cut defined in the text). Column (5) allows both the intercept and the age gradient to differ in lower-density regions by including an indicator for (Density < 95 pct.) and its interaction with age.

Across specifications, the estimated age gradient is very similar to our baseline: the demand for skilled labor declines with technology age at roughly the same rate in high- and lower-density regions. The interaction term in column (5) is close to zero and statistically insignificant, indicating no meaningful difference in the level of skill intensity and the rate at which this skill intensity falls with technology age across regions of different densities.

C Role of γ for model-based decomposition

This section explores the robustness of the model-based decomposition to using different values for the elasticity γ, ranging from 1.2 to 2. Table A5 reports our findings. The table presents the observed change in the college premium (column 1), the total change in the relative demand for college workers inferred from the data (column 2), and the contribution of changes in the pace of technology creation (columns 3 and 4).

Table A5: Decomposition of changes in the college premium, 1970–2020.

	Observed Change (log points)	Inferred shift in relative demand	Contribution from changes in pace of technology	Share explained by changes in the pace of technology
	(1)	(2)	(3)	(4)
Value for γ :				
1.2	29.2	-71.0	34.4	34.4%
1.4	29.2	-68.4	31.5	32.3%
1.6	29.2	-65.7	28.6	30.1%
1.8	29.2	-62.9	25.6	27.8%
2.0	29.2	-60.1	22.6	25.3%

Notes: The table reports the results from model-based decompositions of the college premium using different values for γ. The table presents the observed change in the college premium (column 1), the total inferred change in the relative demand for college workers (column 2), and the contribution of changes in the pace of technology creation (columns 3 and 4).

D Theoretical Derivations and Proofs

This section presents derivations and proofs. We first provide a technical lemma that we use to establish the uniqueness of the equilibrium in our model and to sign comparative statics. We then provide derivations for the baseline model and prove Propositions 1 and 2. We then provide derivations for the spatial model and prove Propositions 3 and 4. Finally, we derive the equilibrium conditions for the model with demographics and prove Propositions 5 and 6.

D.1 Lemmas for uniqueness and signing comparative statics

We first show that the system of equations formed by (1), (2), and (3) has a unique solution.

Lemma A1 (Existence and uniqueness)**.** *Consider a system of equations of the form*

$$1 \; = \; \int_0^\infty f(u)\; c\big(\alpha(u),\, W_h, W_\ell\big)^{1-\sigma}\, du, \tag{A4}$$

$$W_h\, h \;=\; Y \int_0^\infty f(u)\, c\big(\alpha(u), W_h, W_\ell\big)^{1-\sigma}\, \alpha(u) \left(\frac{W_h}{c\big(\alpha(u), W_h, W_\ell\big)}\right)^{1-\gamma} du, \tag{A5}$$

$$W_\ell\, \ell \;=\; Y \int_0^\infty f(u)\, c\big(\alpha(u), W_h, W_\ell\big)^{1-\sigma} \left[1-\alpha(u)\right] \left(\frac{W_\ell}{c\big(\alpha(u), W_h, W_\ell\big)}\right)^{1-\gamma} du, \tag{A6}$$

where h is an L^1 function in $[0,\infty)$ (i.e., an integrable function with finite integral). This system admits a unique solution in W_h, W_ℓ, Y.

Proof. Dividing (A5) by (A6) yields

$$\frac{\int_0^\infty f(u)\, c\big(\alpha(u), W_h, W_\ell\big)^{1-\sigma}\, \alpha(u) \left(\frac{W_h}{c(\alpha(u), W_h, W_\ell)}\right)^{1-\gamma} du}{\int_0^\infty f(u)\, c\big(\alpha(u), W_h, W_\ell\big)^{1-\sigma} \left[1-\alpha(u)\right] \left(\frac{W_\ell}{c(\alpha(u), W_h, W_\ell)}\right)^{1-\gamma} du} \;=\; \frac{W_h\, h}{W_\ell\, \ell}.$$

Using the fact that c is homogeneous of degree 1, we can rewrite this equation in terms of $\omega = W_h/W_\ell$ as

$$\frac{\displaystyle\int_0^\infty f(u)\, c\big(\alpha(u), \omega, 1\big)^{\gamma-\sigma}\, \alpha(u)\, du}{\displaystyle\int_0^\infty f(u)\, c\big(\alpha(u), \omega, 1\big)^{\gamma-\sigma} \left[1-\alpha(u)\right] du}\; \omega^{-\gamma} \;=\; \frac{h}{\ell}. \tag{A7}$$

We first show that this equation has a unique solution ω^*. Write this equation in logs as $g(\ln\omega) = \ln(h/\ell)$, where $\ln\omega \in (-\infty, \infty)$. The derivative of g with respect to $\ln\omega$ is

$$g'(\ln\omega) \;=\; -\underbrace{\Big[(\sigma-\gamma)\,(A_h + A_\ell - 1) + \gamma\Big]}_{\sigma_{agg}},$$

where σ_{agg} is the aggregate elasticity of substitution, as in Oberfield and Raval (2021). This elasticity depends on σ, γ,

$$A_h \;=\; \int_0^\infty \chi_h(u)\, s_h(u)\, du, \text{ with: } \chi_h(u) = \frac{h(u)}{\int_0^\infty h(\tilde{u})d\tilde{u}},\; s_h(u) = \frac{\omega\; h(u)}{\omega\; h(u) \;+\; \ell(u)},$$

and

$$A_\ell \;=\; \int_0^\infty \chi_\ell(u)\, s_\ell(u)\, du, \text{ with: } \chi_\ell(u) = \frac{\ell(u)}{\int_0^\infty \ell(\tilde{u}) d\tilde{u}},\, s_\ell(u) = \frac{\ell(u)}{\omega\, h(u) \;+\; \ell(u)}.$$

In these definitions,

$$h(u) \;=\; f(u)\, c\big(\alpha(u), \omega, 1\big)^{\gamma-\sigma}\; \alpha(u)\, \omega^{-\gamma}$$

and

$$\ell(u) \;= f(u)\, c\big(\alpha(u), \omega, 1\big)^{\gamma-\sigma}\; \big[1 - \alpha(u)\big].$$

The aggregate elasticity of substitution is always positive and is a convex combination of γ and σ. This follows from the fact that

$$2 \geq A_h + A_\ell \geq 1.$$

The left inequality follows from adding up

$$1 = \int_0^\infty \chi_h(u)\, du \geq \int_0^\infty \chi_h(u)\, s_h(u)\, du$$

and

$$1 = \int_0^\infty \chi_\ell(u)\, du \geq \int_0^\infty \chi_\ell(u)\, s_\ell(u)\, du.$$

The right inequality follows from an application of Cauchy-Schwarz, which implies

$$\left(\int_0^\infty \omega\, h(u)\, \frac{\omega\, h(u)}{\omega\, h(u) \;+\; \ell(u)}\, du\right) \left(\int_0^\infty [\omega\, h(u) \;+\; \ell(u)]\, du\right) \geq \left(\int_0^\infty \omega\, h(u)\, du\right)^2$$

or, equivalently,

$$A_h \;=\; \int_0^\infty \chi_h(u)\, s_h(u)\, du \;\geq\; \frac{\int_0^\infty \omega\, h(u)\, du}{\int_0^\infty [\omega\, h(u) \;+\; \ell(u)]\, du} \;=\; s_h,$$

with s_h the share of skilled labor in GDP. Likewise,

$$A_\ell \;=\; \int_0^\infty \chi_\ell(u)\, s_\ell(u)\, du \;\geq\; \frac{\int_0^\infty \ell(u)\, du}{\int_0^\infty [\omega\, h(u) \;+\; \ell(u)]\, du} \;=\; s_\ell,$$

with s_ℓ the share of low-skill labor in GDP. Adding these yields $A_h + A_\ell \geq 1$.

Because $A_h + A_\ell$ is in $[1, 2]$, the aggregate elasticity is always between σ (when $A_h + A_\ell = 2$) and γ (when $A_h + A_\ell = 1$).

We now show that this implies the existence of a unique solution ω^*.

Our strategy is to show that $g(x) \to \infty$ as $x \to -\infty$ and $g(x) \to -\infty$ as $x \to \infty$. The existence of a unique solution then follows from the intermediate value theorem and the fact that g is strictly decreasing.

Take $x = \ln \omega$ and $x_0 = \ln \omega_0$ with $x < x_0$. We have:

$$g(x) - g(x_0) \;=\; \int_x^{x_0} [-g'(s)]\, ds \geq \min\{\gamma, \sigma\}\,(x_0 - x).$$

This implies $g(x) \to \infty$ as $x \to -\infty$. Now take $x = \ln \omega$ and $x_0 = \ln \omega_0$ with $x > x_0$. We have:

$$g(x) - g(x_0) \;=\; \int_{x_0}^{x} g'(s)\, ds \leq -\min\{\gamma, \sigma\}\,(x - x_0).$$

This implies $g(x) \to -\infty$ as $x \to \infty$, as claimed.

The above argument shows that there exists a unique ω^* that solves (A7). We now turn to the level of wages and output. We can write $W_h = \omega\, W_\ell$. Plug this in (A4) to obtain

$$W_\ell^* \;=\; \left(\int_0^\infty f(u)\; c\big(\alpha(u),\, \omega^*, 1\big)^{1-\sigma}\, du\right)^{\frac{1}{\sigma-1}}.$$

Finally, the equilibrium level of output can be obtained as

$$Y^* = W_h^*\, h + W_\ell^*\, \ell \;=\; (\omega^* h + 1)\left(\int_0^\infty f(u)\; c\big(\alpha(u),\, \omega^*, 1\big)^{1-\sigma}\, du\right)^{\frac{1}{\sigma-1}}.$$

□

Lemma A2 (Relative demand curve is downward slopping)**.** *In equation(A7), the equilibrium skill* *premium ω^* (i) rises with any change in $f(u)$ that increase the ratio*

$$\frac{\int_0^\infty f(u)\, c\big(\alpha(u),\omega,1\big)^{\gamma-\sigma}\ \alpha(u)\, du}{\int_0^\infty f(u)\, c\big(\alpha(u),\omega,1\big)^{\gamma-\sigma}\ [1-\alpha(u)]\, du}$$

(holding wages constant), and (ii) decreases in h/ℓ. In particular, any shift in $f(u)$ to $f^{new}(u)$ with

$$\frac{f^{new}(u)}{f(u)} > \frac{f^{new}(u')}{f(u')} \quad \textit{for all} \quad u' > u, \tag{A8}$$

raises the skill premium in (A7).

Proof. Statements (i) and (ii) follow from the fact that the left side of (A7) is decreasing and crosses the right side at a unique point ω^*.

For the second part of the lemma, it suffices to show that

$$\frac{\int_0^\infty f^{new}(u)\, c\big(\alpha(u),\omega,1\big)^{\gamma-\sigma}\ \alpha(u)\, du}{\int_0^\infty f^{new}(u)\, c\big(\alpha(u),\omega,1\big)^{\gamma-\sigma}\ [1-\alpha(u)]\, du} > \frac{\int_0^\infty f(u)\, c\big(\alpha(u),\omega,1\big)^{\gamma-\sigma}\ \alpha(u)\, du}{\int_0^\infty f(u)\, c\big(\alpha(u),\omega,1\big)^{\gamma-\sigma}\ [1-\alpha(u)]\, du}$$

To prove this inequality, observe that, for all $u' \neq u$, we have

$$\left[\frac{f^{new}(u)}{f(u)} - \frac{f^{new}(u')}{f(u')}\right] \left[\frac{1-\alpha(u')}{\alpha(u')} - \frac{1-\alpha(u)}{\alpha(u)}\right] K(u)\ \ K(u') > 0,$$

where

$$K(u) \ = \ f(u)\, c(\alpha(u),\omega,1)^{\gamma-\sigma}\ \alpha(u)$$

is an L^1 Kernel function. Integrating over $u \in [0,\infty)$ and $u' \in [0,\infty)$ yields

$$\int_{u=0}^\infty \int_{u'=0}^\infty \left[\frac{f^{new}(u)}{f(u)} - \frac{f^{new}(u')}{f(u')}\right] \left[\frac{1-\alpha(u')}{\alpha(u')} - \frac{1-\alpha(u)}{\alpha(u)}\right] K(u)\ \ K(u')\ du\ du' > 0,$$

since points with $u' = u$ contribute zero to the integral. Expanding terms yields

$$
\begin{aligned}
-\int_{u=0}^{\infty} \frac{f^{new}(u)}{f(u)} \frac{1-\alpha(u)}{\alpha(u)} K(u)\, du \int_{u'=0}^{\infty} K(u')\, du' \\
+ \int_{u=0}^{\infty} \frac{1-\alpha(u)}{\alpha(u)} K(u)\, du \int_{u'=0}^{\infty} \frac{f^{new}(u')}{f(u')} K(u')\, du' \\
+ \int_{u=0}^{\infty} \frac{f^{new}(u)}{f(u)} K(u)\, du \int_{u'=0}^{\infty} \frac{1-\alpha(u')}{\alpha(u')} K(u')\, du' \\
- \int_{u=0}^{\infty} K(u)\, du \int_{u'=0}^{\infty} \frac{f^{new}(u')}{f(u')} \frac{1-\alpha(u')}{\alpha(u')} K(u')\, du' > 0
\end{aligned}
$$

The negative terms are equal, but use different labels for the variable of integration. The positive terms are also equal, but the variable of integration switches labels. Pooling them and rearranging yields

$$
2\int_0^{\infty} \frac{f^{new}(u)}{f(u)} K(u)\, du \int_0^{\infty} \frac{1-\alpha(u)}{\alpha(u)} K(u)\, du > 2\int_0^{\infty} \frac{f^{new}(u)}{f(u)} \frac{1-\alpha(u)}{\alpha(u)} K(u)\, du \int_0^{\infty} K(u)\, du,
$$

or equivalently

$$
\frac{\int_0^{\infty} \frac{f^{new}(u)}{f(u)} K(u)\, du}{\int_0^{\infty} \frac{f^{new}(u)}{f(u)} \frac{1-\alpha(u)}{\alpha(u)} K(u)\, du} > \frac{\int_0^{\infty} K(u)\, du}{\int_0^{\infty} \frac{1-\alpha(u)}{\alpha(u)} K(u)\, du}
$$

Plugging back in the Kernel, we obtain

$$
\frac{\int_0^{\infty} f^{new}(u)\, c(\alpha(u), \omega, 1)^{\gamma-\sigma}\, \alpha(u)\, du}{\int_0^{\infty} f^{new}(u)\, c(\alpha(u), \omega, 1)^{\gamma-\sigma}\, [1-\alpha(u)]\, du} > \frac{\int_0^{\infty} f(u)\, c(\alpha(u), \omega, 1)^{\gamma-\sigma}\, \alpha(u)\, du}{\int_0^{\infty} f(u)\, c(\alpha(u), \omega, 1)^{\gamma-\sigma}\, [1-\alpha(u)]\, du},
$$

as wanted. □

D.2 Proofs for the Baseline Economy

This section derives equilibrium conditions for the baseline economy and proves Propositions 1 and 2.

Preliminaries: we provide a derivation for equations (1), (2), and (3).

Let's normalize the price of the final good to 1. Technologies of age u are priced at a marginal cost, which implies

$$p_t(u) \;=\; \frac{c\big(\alpha(u),\, W_{h,t}, W_{\ell,t}\big)}{A(t-u)\, z(u)}.$$

The demand for age u technologies at time t is then

$$y_t(u) \;=\; Y_t \left(\frac{c\big(\alpha(u),\, W_{h,t}, W_{\ell,t}\big)}{A(t-u)\, z(u)} \right)^{-\sigma}.$$

This also implies that payments per technology are

$$p_t(u)\; y_t(u) \;=\; Y_t \left(\frac{c\big(\alpha(u),\, W_{h,t}, W_{\ell,t}\big)}{A(t-u)\, z(u)} \right)^{1-\sigma}.$$

Adding payments across all technologies yields

$$Y_t = \int_0^\infty m_t(u)\; p_t(u)\; y_t(u)\; du \;=\; Y_t \int_0^\infty m_t(u) \left(\frac{c\big(\alpha(u),\, W_{h,t}, W_{\ell,t}\big)}{A(t-u)\, z(u)} \right)^{1-\sigma} du.$$

Canceling Y_t on both sides yields equation (1).

We now compute payments to skilled workers. Each technology of age u hires $h_t(u)$ workers and pays them a fraction

$$s_{h,t}(u) \;=\; \alpha(u) \left(\frac{W_{h,t}}{c\big(\alpha(u), W_{h,t}, W_{\ell,t}\big)} \right)^{1-\gamma}$$

of the income generated. Because the market for skilled workers clears, we must have that total payments to skilled labor equal the sum of payments across technologies, or

$$\begin{aligned} W_{h,t}\; h \;&=\; \int_0^\infty m_t(u)\; p_t(u)\; y_t(u)\; s_{h,t}(u)\; du \\ &= Y_t \int_0^\infty m_t(u) \left(\frac{c\big(\alpha(u),\, W_{h,t}, W_{\ell,t}\big)}{A(t-u)\, z(u)} \right)^{1-\sigma} \alpha(u) \left(\frac{W_{h,t}}{c\big(\alpha(u), W_{h,t}, W_{\ell,t}\big)} \right)^{1-\gamma} du, \end{aligned}$$

which gives equation (2).

For low-skill workers, each technology of age u hires $\ell_t(u)$ workers and pays them a fraction

$$s_{\ell,t}(u) \;=\; [1-\alpha(u)] \left(\frac{W_{\ell,t}}{c\big(\alpha(u), W_{h,t}, W_{\ell,t}\big)}\right)^{1-\gamma}$$

of the income generated. Because the market for low-skill workers clears, we must have that total payments to skilled labor equal the sum of payments across technologies, or

$$\begin{aligned} W_{\ell,t}\;\ell \;&=\; \int_0^\infty m_t(u)\; p_t(u)\; y_t(u)\; s_{\ell,t}(u)\; du \\ &= Y_t \int_0^\infty m_t(u) \left(\frac{c\big(\alpha(u),\, W_{h,t}, W_{\ell,t}\big)}{A(t-u)\, z(u)}\right)^{1-\sigma} [1-\alpha(u)] \left(\frac{W_{\ell,t}}{c\big(\alpha(u), W_{h,t}, W_{\ell,t}\big)}\right)^{1-\gamma} du, \end{aligned}$$

which gives equation (3).

Proofs of Propositions: we now prove Propositions 1 and 2.

Proof of Proposition 1. Suppose $m_t(u) = m$. Substitute $w_{h,t} = W_{h,t}/A(t)$, $w_{\ell,t} = W_{\ell,t}/A(t)$ and $y(t) = Y(t)/A(t)$ in (1), (2), and (3) to obtain:

$$\begin{aligned} 1 \;&=\; \int_0^\infty m \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u),\, w_{h,t}, w_{\ell,t}\big)^{1-\sigma} du, \\ w_{h,t}\; h \;&= y_t \int_0^\infty m \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u), w_{h,t}, w_{\ell,t}\big)^{1-\sigma} \alpha(u) \left(\frac{w_{h,t}}{c\big(\alpha(u), w_{h,t}, w_{\ell,t}\big)}\right)^{1-\gamma} du \\ w_{\ell,t}\; \ell \;&= y_t \int_0^\infty m \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u), w_{h,t}, w_{\ell,t}\big)^{1-\sigma} \big[1-\alpha(u)\big] \left(\frac{w_{\ell,t}}{c\big(\alpha(u), w_{h,t}, w_{\ell,t}\big)}\right)^{1-\gamma} du. \end{aligned}$$

The function $f(u) = m\; (z(u)/A(u))^{\sigma-1}$ is L^1 over $[0,\infty)$, since we assumed $g > g_{\text{lim}}$. Lemma A1 then implies there exists a unique solution to the above system with $w_{h,t} = w_h^*$, $w_{\ell,t} = w_\ell^*$, and $y_t = y^*$. In terms of the initial aggregate variables, this implies that $W_{h,t} = w_h^*\; A(t)$, $W_{\ell,t} = w_\ell^*\; A(t)$, and $Y_t = y^*\; A(t)$, which means all aggregates frow at a rate of g, establishing the fact that the economy is in a balanced growth path.

From the derivations in Lemma A1, the college premium is given by

$$\frac{\displaystyle\int_0^\infty m \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u), \omega, 1\big)^{\gamma-\sigma}\ \alpha(u)\, du}{\displaystyle\int_0^\infty m \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u), \omega, 1\big)^{\gamma-\sigma}\ [1-\alpha(u)]\, du}\ \omega^{-\gamma} \ = \ \frac{h}{\ell}.$$

The solution to this equation is independent of m, establishing the fact that the balanced-growth path skill premium is constant and independent of m. □

Proof of Proposition 2. Suppose the increase to m' lasts for at least T periods, for some $T > 0$.

For $t \in (t_0, t_0+T)$, the skill premium is given by equation (A7) with $f^{new}(u) = m'\ (z(u)/A(u))^{\sigma-1}$ for $u < t$ and $f^{new}(u) = m\ (z(u)/A(u))^{\sigma-1}$ for $u > u'$. Instead, at t_0, the skill premium is given by equation (A7) with$f(u) = m\ (z(u)/A(u))^{\sigma-1}$.

Lemma A2 implies that the skill premium rises for all $t \in (t_0, t_0 + T)$ because $f^{new}(u)$ and $f(u)$ satisfy inequality (A8).

The fact that the skill premium reverts to its initial level follows from the fact that the balanced-growth path level of the skill premium is independent of m. □

D.3 Proofs for the Spatial Economy

This section derives equilibrium conditions and proves Propositions 3 and 4.

Preliminaries: we first derive the equilibrium conditions for the spatial model.

Let's normalize the price of the final good to 1 in each location (this can be done because these locations are in autarky). Technologies of age u are priced at a marginal cost, which implies

$$p_t(u,d) \ = \ \frac{c\big(\alpha(u),\ W_{h,t}(d), W_{\ell,t}(d)\big)}{A(t-u)\, z(u)}.$$

The demand for age u technologies at time t is then

$$y_t(u,d) \ = \ Y_t(d) \left(\frac{c\big(\alpha(u),\ W_{h,t}(d), W_{\ell,t}(d)\big)}{A(t-u)\, z(u)}\right)^{-\sigma}.$$

This also implies that payments per technology are

$$p_t(u,d)\; y_t(u,d) \;=\; Y_t(d) \left(\frac{c\big(\alpha(u),\, W_{h,t}(d), W_{\ell,t}(d)\big)}{A(t-u)\, z(u)} \right)^{1-\sigma}.$$

Adding payments across all technologies yields

$$\begin{aligned} Y_t(d) &= \int_0^\infty p(u,d)\; m_t(u)\; p_t(u,d)\; y_t(u,d)\; du \\ &= Y_t(d) \int_0^\infty p(u,d)\; m_t(u) \left(\frac{c\big(\alpha(u),\, W_{h,t}(d), W_{\ell,t}(d)\big)}{A(t-u)\, z(u)} \right)^{1-\sigma} du. \end{aligned}$$

Canceling $Y_t(d)$ on both sides yields the spatial analog of equation (1):

$$1 \;=\; \int_0^\infty p(u,d)\; m_t(u) \left(\frac{c\big(\alpha(u),\, W_{h,t}(d), W_{\ell,t}(d)\big)}{A(t-u)\, z(u)} \right)^{1-\sigma} du. \tag{A9}$$

We now compute payments to skilled workers. Each technology of age u hires $h_t(u,d)$ workers and pays them a fraction

$$s_{h,t}(u,d) \;=\; \alpha(u) \left(\frac{W_{h,t}(d)}{c\big(\alpha(u), W_{h,t}(d), W_{\ell,t}(d)\big)} \right)^{1-\gamma}$$

of the income generated. Because the market for skilled workers clears, we must have that total payments to skilled labor equal the sum of payments across technologies, or

$$\begin{aligned} W_{h,t}(d)\; h(d) \;&=\; \int_0^\infty p(u,d)\; m_t(u)\; p_t(u,d)\; y_t(u,d)\; s_{h,t}(u,d)\; du \\ &= Y_t(d) \int_0^\infty p(u,d)\; m_t(u) \left(\frac{c\big(\alpha(u),\, W_{h,t}(d), W_{\ell,t}(d)\big)}{A(t-u)\, z(u)} \right)^{1-\sigma} \alpha(u) \left(\frac{W_{h,t}(d)}{c\big(\alpha(u), W_{h,t}(d), W_{\ell,t}(d)\big)} \right)^{1-\gamma} du, \end{aligned}$$

which gives an analog to equation (2):

$$W_{h,t}(d)\; h(d) \;=\; Y_t(d) \int_0^\infty p(u,d)\; m_t(u) \left(\frac{c\big(\alpha(u),\, W_{h,t}(d), W_{\ell,t}(d)\big)}{A(t-u)\, z(u)} \right)^{1-\sigma} \alpha(u) \left(\frac{W_{h,t}(d)}{c\big(\alpha(u), W_{h,t}(d), W_{\ell,t}(d)\big)} \right)^{1-\gamma} du. \tag{A10}$$

For low-skill workers, each technology of age u hires $\ell_t(u,d)$ workers and pays them a fraction

$$s_{\ell,t}(u,d) \;=\; [1-\alpha(u)]\left(\frac{W_{\ell,t}(d)}{c\big(\alpha(u),W_{h,t}(d),W_{\ell,t}(d)\big)}\right)^{1-\gamma}$$

of the income generated. Because the market for low-skill workers clears, we must have that total payments to skilled labor equal the sum of payments across technologies, or

$$\begin{aligned} W_{\ell,t}(d)\,\ell(d) \;&=\; \int_0^\infty p(u,d)\,m_t(u)\,p_t(u,d)\,y_t(u,d)\,s_{\ell,t}(u,d)\,du \\ &= Y_t(d)\int_0^\infty p(u,d)\,m_t(u)\left(\frac{c\big(\alpha(u),\,W_{h,t}(d),W_{\ell,t}(d)\big)}{A(t-u)\,z(u)}\right)^{1-\sigma}[1-\alpha(u)]\left(\frac{W_{\ell,t}(d)}{c\big(\alpha(u),W_{h,t}(d),W_{\ell,t}(d)\big)}\right)^{1-\gamma}du, \end{aligned}$$

which gives an analog to equation (3):

$$W_{\ell,t}(d)\,\ell(d) \;=\; Y_t(d)\int_0^\infty p(u,d)\,m_t(u)\left(\frac{c\big(\alpha(u),\,W_{h,t}(d),W_{\ell,t}(d)\big)}{A(t-u)\,z(u)}\right)^{1-\sigma}[1-\alpha(u)]\left(\frac{W_{\ell,t}(d)}{c\big(\alpha(u),W_{h,t}(d),W_{\ell,t}(d)\big)}\right)^{1-\gamma}du. \tag{A11}$$

Proofs of Propositions: we now prove Propositions 3 and 4.

Proof of Proposition 3. Suppose $m_t(u) = m$. Substitute $w_{h,t}(d) = W_{h,t}(d)/A(t), w_{\ell,t}(d) = W_{\ell,t}(d)/A(t)$ and $y_t(d) = Y_t(d)/A(t)$ in (A9), (A10), and (A11) to obtain:

$$\begin{aligned} 1 \;&=\; \int_0^\infty p(u,d)\;m\;\left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u),\,w_{h,t}(d),w_{\ell,t}(d)\big)^{1-\sigma}\,du, \\ w_{h,t}(d)\,h(d) \;&=\; y_t(d)\int_0^\infty p(u,d)\;m\;\left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u),w_{h,t}(d),w_{\ell,t}(d)\big)^{1-\sigma} \\ &\qquad\qquad \alpha(u)\left(\frac{w_{h,t}(d)}{c\big(\alpha(u),w_{h,t}(d),w_{\ell,t}(d)\big)}\right)^{1-\gamma}du \\ w_{\ell,t}(d)\,\ell(d) \;&=\; y_t(d)\int_0^\infty p(u,d)\;m\;\left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u),w_{h,t}(d),w_{\ell,t}(d)\big)^{1-\sigma} \\ &\qquad\qquad \big[1-\alpha(u)\big]\left(\frac{w_{\ell,t}(d)}{c\big(\alpha(u),w_{h,t}(d),w_{\ell,t}(d)\big)}\right)^{1-\gamma}du. \end{aligned}$$

The function $f(u) = p(u,d)\;m\;(z(u)/A(u))^{\sigma-1}$ is L^1 over $[0,\infty)$, since we assumed $g > g_{\lim}$.

Lemma A1 then implies there exists a unique solution to the above system with $w_{h,t}(d) = w_h^*(d)$, $w_{\ell,t}(d) = w_\ell^*(d)$, and $y_t(d) = y^*(d)$. In terms of the initial aggregate variables, this implies that $W_{h,t}(d) = w_h^*(d)\ A(t)$, $W_{\ell,t}(d) = w_\ell^*(d)\ A(t)$, and $Y_t(d) = y^*(d)\ A(t)$, which means all aggregates frow at a rate of g, establishing the fact that the economy is in a balanced growth path.

From the derivations in Lemma A1, the college balanced-growth path college premium in a location of density d is pinned down by

$$\frac{\int_0^\infty p(u,d)\ m\left(\frac{z(u)}{A(u)}\right)^{\sigma-1}\ c\big(\alpha(u),\omega(d),1\big)^{\gamma-\sigma}\ \alpha(u)\,du}{\int_0^\infty p(u,d)\ m\left(\frac{z(u)}{A(u)}\right)^{\sigma-1}\ c\big(\alpha(u),\omega(d),1\big)^{\gamma-\sigma}\ [1-\alpha(u)]\,du}\ \omega(d)^{-\gamma} = \frac{h(d)}{\ell(d)}.$$

The solution to this equation is independent of m, establishing the fact that the balanced-growth path skill premium is constant and independent of m.

Moreover, suppose $h(d)/\ell(d)$ is constant. Then Lemma A2 implies that $\omega(d)$ is increasing in d. This is because the log-submodularity of $p(u,d)$ implies that, for $d' > d$ and $u' > u$, we have

$$\frac{p(u,d')}{p(u,d)} > \frac{p(u',d')}{p(u',d)},$$

which implies that inequality (A8) holds. □

Proof of Proposition 4. Proposition 2 implies that the skill premium rises along the transition in all locations. We now show that early on, the rise is more pronounced in dense locations.

Let's suppose $m(b)$ rises from m to m' for $b \in [t_0, \bar{t})$, where $\bar{t}$ can be infinite if the shock is permanent. Take two cities with $d' > d$, and assume $h(d)/\ell(d)$ is constant across d. We show that as t approaches t_0,

$$\ln \omega_t(d') - \ln \omega_{t_0}(d') > \ln \omega_t(d) - \ln \omega_{t_0}(d).$$

This is equivalent to showing that

$$\frac{\partial \ln \omega_t(d)}{\partial t} \text{ increases in } d \text{ at } t = t_0.$$

Take any $t \in (t_0, \bar{t})$. The skill premium at t in location d solves

$$\frac{\displaystyle\int_0^\infty p(u,d)\,\big((m'-m)\,\mathbb{1}[u \le t-t_0]+m\big)\left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u),\omega_t(d),1\big)^{\gamma-\sigma}\,\alpha(u)\,du}{\displaystyle\int_0^\infty p(u,d)\,\big((m'-m)\,\mathbb{1}[u \le t-t_0]+m\big)\left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u),\omega(d),1\big)^{\gamma-\sigma}\,[1-\alpha(u)]\,du}\;\omega_t(d)^{-\gamma} \;=\; \tfrac{h}{\ell}.$$

Log-differentiate this expression with respect to t around $t = t_0$ to obtain:

$$\frac{\partial \ln \omega_t(d)}{\partial t} \;=\; \frac{m'-m}{m}\;\frac{\chi_h(0,d)-\chi_\ell(0,d)}{\sigma_{agg}(d)},$$

where, as in the proof of Lemma A1,

$$\sigma_{agg}(d) \;=\; (\sigma-\gamma)\,\big(A_h(d)+A_\ell(d)-1\big)+\gamma,$$

with $\sigma_{agg} > 0$ the aggregate elasticity of substitution, as in Oberfield and Raval (2021). This elasticity depends on σ, γ,

$$A_h(d) \;=\; \int_0^\infty \chi_h(u,d)\,s_h(u,d)\,du, \text{ with: } \chi_h(u,d) = \frac{h(u,d)}{\int_0^\infty h(\tilde{u},d)d\tilde{u}},\; s_h(u,d) = \frac{\omega(d)\,h(u,d)}{\omega(d)\,h(u,d)+\ell(u,d)},$$

and

$$A_\ell(d) \;=\; \int_0^\infty \chi_\ell(u,d)\,s_\ell(u,d)\,du, \text{ with: } \chi_\ell(u,d) = \frac{\ell(u,d)}{\int_0^\infty \ell(\tilde{u},d)d\tilde{u}},\; s_\ell(u,d) = \frac{\ell(u,d)}{\omega(d)\,h(u,d)+\ell(u,d)}.$$

In these definitions,

$$h(u,d) \;=\; p(u,d)\,m\left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u),\omega(d),1\big)^{\gamma-\sigma}\,\alpha(u)\,\omega(d)^{-\gamma}$$

and

$$\ell(u) \;= p(u,d)\,m\left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\big(\alpha(u),\omega(d),1\big)^{\gamma-\sigma}\,\big[1-\alpha(u)\big].$$

Finally, $\chi_h(0,d)$ gives the share of skilled worker employment in technologies of age $u = 0$ (as a fraction of all skilled employment), and $\chi_\ell(0,d)$ gives the share of low-skill worker employment in technologies of age $u = 0$ (as a fraction of all low-skill employment).

Note that $\chi_h(0,d) > \chi_\ell(0,d)$ (from the fact that $\alpha(u)$ decreases). This implies, in line with Proposition 2, that $\frac{\partial \ln \omega_t(d)}{\partial t} > 0$ early on in the transition.

Moreover, these derivations show that $\frac{\partial \ln \omega_t(d)}{\partial t}$ increases in d if and only if

$$\frac{\chi_h(0,d) - \chi_\ell(0,d)}{\sigma_{agg}(d)}$$

increases in d, as claimed in the proposition.

This holds, in particular, if $\sigma = \gamma$ and $\alpha(0) > \bar{\alpha}$. To see this, take the limit case with $\alpha(0) = 1$ (and $\sigma = \gamma$). Then

$$\frac{\chi_h(0,d) - \chi_\ell(0,d)}{\sigma_{agg}(d)} = \frac{\chi_h(0,d)}{\sigma} = \frac{1}{\sigma} \frac{1}{\int_0^\infty \frac{p(u,d)}{p(0,d)} \left(\frac{z(u)/z(0)}{A(u)/A(0)}\right)^{\sigma-1} \frac{\alpha(u)}{\alpha(0)}\, du}.$$

The right side increases in d because the log-submodularity of $p(u,d)$ implies that $p(u,d)/p(0,d)$ decreases in d. By continuity, this is also true for $\alpha(0) > \bar{\alpha}$. □

D.4 Proofs for the Economy with Demographics

This section derives equilibrium conditions for the economy with demographics and proves Propositions 5 and 6.

Preliminaries: We first derive the equilibrium conditions for the demographic model. We focus on an equilibrium with full specialization, where (i) skilled workers of age u_p are assigned to a unique technology of age $u = \mathcal{U}_{h,t}(u_p)$ and paid a wage $W_{h,t}(u_p)$, (ii) low-skill workers of age u_p are assigned to a unique technology of age $u = \mathcal{U}_{\ell,t}(u_p)$ and paid a wage $W_{\ell,t}(u_p)$, (iii) both assignment functions are strictly increasing, so that young workers specialize in recently introduced technologies, and (iv) markets clear. Costinot and Vogel (2010) show that the equilibrium in this class of assignment models always features these properties and can be expressed in terms of a system of differential equations with boundary conditions.

To characterize the equilibrium of our economy, we focus on the inverse mapping from technologies to workers. This is equivalent to the original formulation but makes the math more transparent.

Denote by $\mathcal{M}_{h,t}(u)$ and $\mathcal{M}_{h,t}(u)$ the ages of skilled and low-skill workers producing with technologies of age u, respectively. Denote also by $\mathcal{P}_{h,t}(u)$ and $\mathcal{P}_{\ell,t}(u)$ the wage rate per efficiency unit of skilled and low-skill labor in technologies of age u.

The wage of skilled workers of age u_p is

$$W_{h,t}(u_p) \;=\; \max_u \mathcal{P}_{h,t}(u)\; a(u,u_p)\; e^{\beta_h u_p}.$$

Because workers sort into the technology that maximizes their wage, their choice of $u = \mathcal{U}_{h,t}(u)$ satisfies the necessary first-order condition

$$\frac{\mathcal{P}'_{h,t}\big(\mathcal{U}_{h,t}(u)\big)}{\mathcal{P}_{h,t}\big(\mathcal{U}_{h,t}(u)\big)} = -\frac{\dfrac{\partial a(\mathcal{U}_{h,t}(u_p),u_p)}{\partial u}}{a(\mathcal{U}_{h,t}(u_p),u_p)}.$$

This can be written in terms of $u = \mathcal{U}_{h,t}(u)$ as

$$\frac{\mathcal{P}'_{h,t}(u)}{\mathcal{P}_{h,t}(u)} = -\frac{\dfrac{\partial a\big(u,\mathcal{M}_{h,t}(u)\big)}{\partial u}}{a\big(u,\mathcal{M}_{h,t}(u)\big)}. \tag{A12}$$

The same argument applies to low-skill workers. Their wage is

$$W_{\ell,t}(u_p) \;=\; \max_u \mathcal{P}_{\ell,t}(u)\; a(u,u_p)\; e^{\beta_\ell u_p},$$

and the first-order condition for their optimal job choice yields the differential equation

$$\frac{\mathcal{P}'_{\ell,t}(u)}{\mathcal{P}_{\ell,t}(u)} = -\frac{\dfrac{\partial a\big(u,\mathcal{M}_{\ell,t}(u)\big)}{\partial u}}{a\big(u,\mathcal{M}_{\ell,t}(u)\big)}. \tag{A13}$$

For every $u \in [0,\bar{u}]$, workers of age $\tilde{u}_p \in [0,\mathcal{M}_{h,t}(u)]$ are employed in technologies of age

$\tilde{u} \in [0, u]$. This requires the supply of efficiency units of labor on one side to equal demand:

$$\int_0^{\mathcal{M}_{h,t}(u)} h \; e^{n\,(t-\tilde{u}_p)} \; a(\mathcal{U}_{h,t}(\tilde{u}_p), \tilde{u}_p) \; e^{\beta_h \tilde{u}_p} \; d\tilde{u}_p \;\; = \;\; Y_t \; A(t)^{\sigma-1} \int_0^u m_t(\tilde{u}) \left(\frac{z(\tilde{u})}{A(\tilde{u})}\right)^{\sigma-1} c\Big(\alpha(\tilde{u}), \mathcal{P}_{h,t}(\tilde{u}), \mathcal{P}_{\ell,t}(\tilde{u})\Big)^{\gamma-\sigma} \; \alpha(\tilde{u}) \; \mathcal{P}_{h,t}(\tilde{u})^{-\gamma} d\tilde{u}.$$

Differentiating with respect to u yields

$$\mathcal{M}'_{h,t}(u) \;\; = \;\; \frac{Y_t \; A(t)^{\sigma-1} \; m_t(u) \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\Big(\alpha(u), \mathcal{P}_{h,t}(u), \mathcal{P}_{\ell,t}(u)\Big)^{\gamma-\sigma} \; \alpha(u) \; \mathcal{P}_{h,t}(u)^{-\gamma}}{h \; e^{n\,(t-\mathcal{M}_{h,t}(u))} \; a(u, \mathcal{M}_{h,t}(u)) \; e^{\beta_h \mathcal{M}_{h,t}(u)}}. \tag{A14}$$

This requires the rate of change of the matching function to equal the ratio of the demand for labor to its supply around technology age u. The same calculations for low-skill labor imply

$$\mathcal{M}'_{\ell,t}(u) \;\; = \;\; \frac{Y_t \; A(t)^{\sigma-1} \; m_t(u) \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\Big(\alpha(u), \mathcal{P}_{h,t}(u), \mathcal{P}_{\ell,t}(u)\Big)^{\gamma-\sigma} \; \big[1-\alpha(u)\big] \; \mathcal{P}_{\ell,t}(u)^{-\gamma}}{\ell \; e^{n\,(t-\mathcal{M}_{\ell,t}(u))} \; a(u, \mathcal{M}_{\ell,t}(u)) \; e^{\beta_\ell \mathcal{M}_{\ell,t}(u)}}. \tag{A15}$$

In addition, market-clearing imposes the boundary conditions

$$\mathcal{M}_{h,t}(0) \; = \; 0, \qquad \mathcal{M}_{\ell,t}(0) \; = \; 0, \qquad \mathcal{M}_{h,t}(\bar{u}) = \bar{u}_p, \qquad \mathcal{M}_{\ell,t}(\bar{u}) = \bar{u}_p.$$

Finally, the fact that we normalized the price of the final good to 1 requires

$$1 \; = \; A(t)^{\sigma-1} \int_0^{\bar{u}} m_t(u) \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} c\Big(\alpha(u), \mathcal{P}_{h,t}(u), \mathcal{P}_{\ell,t}(u)\Big)^{1-\sigma} \; du \tag{A16}$$

Equations (A12), (A13), (A14), and (A15) form a system of four differential equations with four boundary conditions and an unknown parameter Y_t pinned down by (A16). We do not provide here an argument for the existence or uniqueness of a solution to this boundary problem, but assume throughout that there is a solution and that it is unique. A solution can be computed numerically using standard solvers, such as Matlab's *bvp4c*, as shown in the replication kit.

Proofs of Propositions: we now prove Propositions 5 and 6.

Proof of Proposition 5. Suppose $m_t(u) = m$. Substitute $p_{h,t}(u) = \mathcal{P}_{h,t}(u)/A(t)$, $p_{\ell,t}(u) = \mathcal{P}_{\ell,t}(u)/A(t)$, and $y_t = Y_t/(A(t)e^{nt})$ in (A12), (A13), (A14), and (A15) to obtain

$$
\begin{aligned}
\frac{p'_{h,t}(u)}{p_{h,t}(u)} &= - \frac{\frac{\partial a\left(u, \mathcal{M}_{h,t}(u)\right)}{\partial u}}{a\left(u, \mathcal{M}_{h,t}(u)\right)}, \\
\frac{p'_{\ell,t}(u)}{p_{\ell,t}(u)} &= - \frac{\frac{\partial a\left(u, \mathcal{M}_{\ell,t}(u)\right)}{\partial u}}{a\left(u, \mathcal{M}_{\ell,t}(u)\right)}, \\
\mathcal{M}'_{h,t}(u) &= \frac{y_t \; m \; \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} \; c\Big(\alpha(u), p_{h,t}(u), p_{\ell,t}(u)\Big)^{\gamma-\sigma} \; \alpha(u) \; p_{h,t}(u)^{-\gamma}}{h \; e^{-n\,\mathcal{M}_{h,t}(u)} \; a(u, \mathcal{M}_{h,t}(u)) \; e^{\beta_h \mathcal{M}_{h,t}(u)}}, \\
\mathcal{M}'_{\ell,t}(u) &= \frac{y_t \; m \; \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} \; c\Big(\alpha(u), p_{h,t}(u), p_{\ell,t}(u)\Big)^{\gamma-\sigma} \; \big[1-\alpha(u)\big] \; p_{\ell,t}(u)^{-\gamma}}{\ell \; e^{-n\,\mathcal{M}_{\ell,t}(u)} \; a(u, \mathcal{M}_{\ell,t}(u)) \; e^{\beta_\ell \mathcal{M}_{\ell,t}(u)}}, \\
1 &= \int_0^{\bar{u}} m \; \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} \; c\Big(\alpha(u), p_{h,t}(u), p_{\ell,t}(u)\Big)^{1-\sigma} \; du
\end{aligned}
$$

The resulting system is stationary, which means that its solution is time independent and given by $y_t = y^*$, $\mathcal{M}_{h,t} = \mathcal{M}^*_h(u)$, $\mathcal{M}_{\ell,t} = \mathcal{M}^*_\ell(u)$, $p_{h,t}(u) = p^*_h(u)$, and $p_{\ell,t}(u) = p^*_\ell(u)$. In terms of aggregates, we have that GDP per capita is $Y_t/e^{nt} = y^* \, A(t)$ and the college premium by age is stable, given by

$$
\omega(u_p) \;=\; \frac{p^*_h(\mathcal{U}_h(u_p)) \; a(\mathcal{U}_h(u_p), u_p)}{p^*_\ell(\mathcal{U}_h(u_p)) \; a(\mathcal{U}_\ell(u_p), u_p)} \; e^{(\beta_h - \beta_\ell) u_p}
$$

□

Proof of Proposition 6. We take $\sigma = \gamma$ and consider the limit case with $\alpha(u) = 1$ for $u \in [0, \underline{u}]$. The results extend to cases with $\alpha(u)$ sufficiently high in $u \in [0, \underline{u}]$ by continuity.

Suppose $m(b)$ rises from m to m' for $b \in [t_0, \bar{t})$, where $\bar{t}$ can be infinite if the shock is permanent. Take a $t \in (t_0, \min\{\bar{t}, t_0 + \underline{u}\})$. We show that for any such t, (i) the matching function for low-skilled workers remains unchanged at its BGP level; and (ii) skilled workers shift toward more recent technologies.

To establish (i), note that equations (A13) and (A15) become

$$\frac{\mathcal{P}'_{\ell,t}(u)}{\mathcal{P}_{\ell,t}(u)} = -\frac{\frac{\partial a(u, \mathcal{M}_{\ell,t}(u))}{\partial u}}{a(u, \mathcal{M}_{\ell,t}(u))}. \tag{A17}$$

and

$$\mathcal{M}'_{\ell,t}(u) = \frac{Y_t\, A(t)^{\sigma-1}\, m_t(u)\, \left(\frac{z(u)}{A(u)}\right)^{\sigma-1} \left[1-\alpha(u)\right] \mathcal{P}_{\ell,t}(u)^{-\gamma}}{\ell\, e^{n\,(t-\mathcal{M}_{\ell,t}(u))}\, a(u, \mathcal{M}_{\ell,t}(u))\, e^{\beta_\ell \mathcal{M}_{\ell,t}(u)}}. \tag{A18}$$

For $t \in (t_0, \min\{\bar{t}, t_0 + \underline{u}\})$ both equations change in time (t) only due to the influence of output Y_t. Changes in $m_t(u)$ have no effect on these equations, since $m_t(u)\,\left[1-\alpha(u)\right] = 0$ early on. This implies that the new values for $\mathcal{M}_{\ell,t}(u)$ and $\mathcal{P}_{\ell,t}(u)$ are given by

$$\mathcal{P}_{\ell,t}(u) = \mathcal{P}^*_{\ell,t}(u) \left(\frac{Y_t}{Y_t^*}\right)^{\frac{1}{\gamma}}, \qquad \mathcal{M}_{\ell,t}(u) = \mathcal{M}^*_{\ell,t},$$

where the $*$s denote the BGP equilibrium allocations and prices.

Our proof for (ii) follows Costinot and Vogel (2010). We first show that skilled workers cannot shift towards strictly older technologies. Suppose this were the case, and let's pick the first time this happens. This means there is an interval of technologies $[u_0, u_1]$ assigned to strictly younger workers, with $\mathcal{M}_{h,t}(u_0) = \mathcal{M}^*_{h,t}(u_0)$ and $\mathcal{M}_{h,t}(u_1) = \mathcal{M}^*_{h,t}(u_1)$. This possibility is depicted in Figure A4, together with the initial BGP matching function for skilled labor.

Because the new matching function cuts the BGP one the first time from above, we have

$$\mathcal{M}'_{h,t}(u_0) < \mathcal{M}^{*\prime}_{h,t}(u_0).$$

and

$$\mathcal{M}'_{h,t}(u_1) > \mathcal{M}^{*\prime}_{h,t}(u_1),$$

Both inequalities combined imply

$$\frac{\mathcal{M}'_{h,t}(u_1)}{\mathcal{M}'_{h,t}(u_0)} > \frac{\mathcal{M}^{*\prime}_{h,t}(u_1)}{\mathcal{M}^{*\prime}_{h,t}(u_0)}.$$

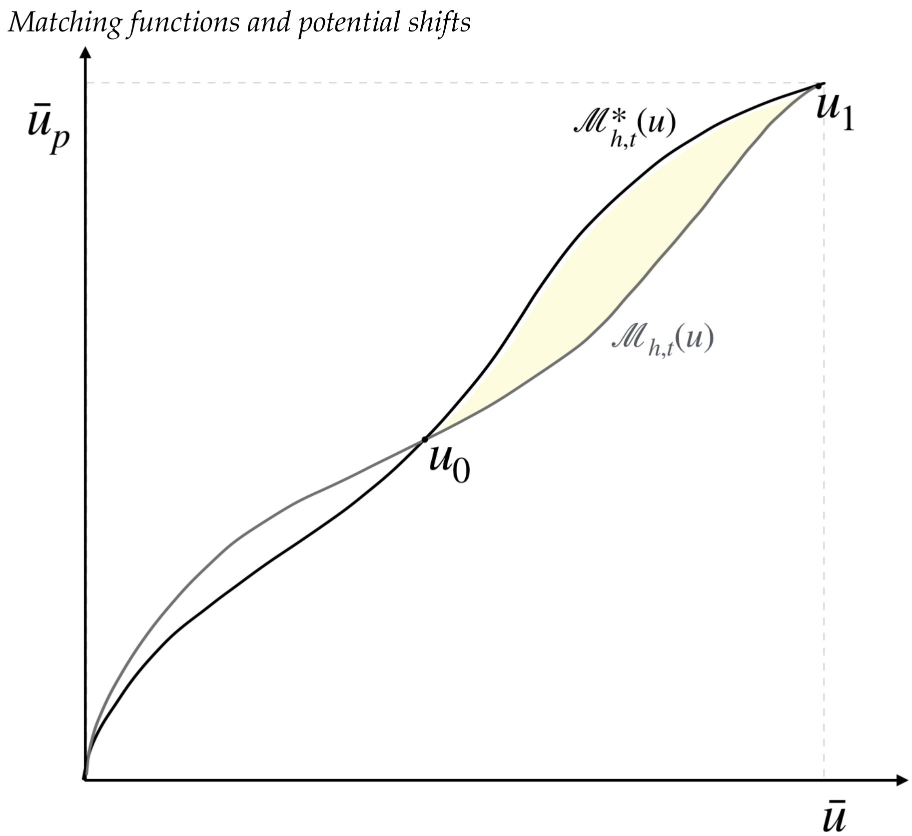

Figure A4: Matching function of technology to workers by age. The figure plots the matching function of technologies to skill workers of age $\mathcal{M}^*_{t,h}(u)$ along the BGP. The figure also illustrates the contradictory possibility that the matching function will shift down for technologies in $[u_0, u_1]$, with workers in this region being matched to older technologies along the transition.

Using equation (A14), this inequality implies

$$\frac{m_t(u_1)}{m_t(u_0)} \left(\frac{\mathcal{P}_{h,t}(u_1)}{\mathcal{P}_{h,t}(u_0)}\right)^{-\gamma} > \frac{m}{m} \left(\frac{\mathcal{P}^*_{h,t}(u_1)}{\mathcal{P}^*_{h,t}(u_0)}\right)^{-\gamma},$$

or, equivalently

$$\left(\frac{\mathcal{P}_{h,t}(u_0)}{\mathcal{P}_{h,t}(u_1)}\right)^{\gamma} > \frac{m_t(u_0)}{m_t(u_1)} \left(\frac{\mathcal{P}^*_{h,t}(u_0)}{\mathcal{P}^*_{h,t}(u_1)}\right)^{\gamma} \geq \left(\frac{\mathcal{P}^*_{h,t}(u_0)}{\mathcal{P}^*_{h,t}(u_1)}\right)^{\gamma}. \tag{A19}$$

This inequality uses the fact that, the increase in $m(b)$ at t_0 (weakly) raises $m_t(u_0)/m_t(u_1)$ for all $t \in (t_0, \min\{\bar{t}, t_0 + \underline{u}\})$.

Let's now turn to equation (A12). Integrating this equation between u_0 and u_1 yields

$$\frac{\mathcal{P}_{h,t}(u_0)}{\mathcal{P}_{h,t}(u_1)} = \exp\left(\int_{u_0}^{u_1} \frac{\frac{\partial a(u, \mathcal{M}_{h,t}(u))}{\partial u}}{a(u, \mathcal{M}_{h,t}(u))}\, du\right) < \exp\left(\int_{u_0}^{u_1} \frac{\frac{\partial a(u, \mathcal{M}^*_{h,t}(u))}{\partial u}}{a(u, \mathcal{M}^*_{h,t}(u))}\, du\right) = \frac{\mathcal{P}^*_{h,t}(u_0)}{\mathcal{P}^*_{h,t}(u_1)}. \quad \text{(A20)}$$

The inequality in the middle follows from the assumption that the new matching function $M_{h,t}(u)$ is below the BGP one $M^*_{h,t}(u)$ for $u \in [u_0, u_1]$ and the fact that the integrand is increasing in u_p (which follows from the log supermodularity of a).

The inequalities (A19) and (A20) contradict each other. Intuitively, if workers are shifting toward u_1, and demand is shifting away from u_1, the relative price of u_0 vis-à-vis u_1 should rise. However, this would push workers toward u_0, contradicting the initial assertion.

The exact same argument can be used to show that the new matching function cannot be equal to the BGP one on $[u_0, u_1]$ and above it for $u > u_1$, or above it for $u < u_0$. In either case, (A19) remains a strict inequality, and (A20) becomes a weak inequality, yet this still implies a contradiction.

The above argument shows that the new matching function must be either identical to the original one or $\mathcal{M}_{h,t}(u)$ must shift up everywhere. We now show that the latter is true. Suppose the matching function for skilled workers remains unchanged for $t \in (t_0, \bar{t}]$. Equation (A12) implies that $\mathcal{P}_{h,t}(u')/\mathcal{P}_{h,t}(u)$ remains unchanged. Dividing (A14) at u' and u then implies that $m_t(u')/m_t(u)$ must be constant for all u', u, which contradicts the fact that this ratio rises for $u' < \bar{t} - t < u$.

We now turn to wages. We first show that the college premium rises for workers of age $\bar{u}_p$. These workers are matched to technologies of age $\bar{u}$. For skilled workers, we have

$$\mathcal{M}'_{ht}(\bar{u}) < \mathcal{M}^{*\prime}_{ht}(\bar{u}),$$

since the matching function shifts up for $t \in (t_0, \min\{\bar{t}, t_0 + \underline{u}\})$. Equation (A14) then implies

$$Y_t\, \mathcal{P}_{ht}(\bar{u})^{-\gamma} < Y^*_t\, \mathcal{P}^*_{ht}(\bar{u})^{-\gamma},$$

or equivalently

$$\mathcal{P}_{h,t}(\bar{u}) \;>\; \mathcal{P}^*_{h,t}(\bar{u}) \left(\frac{Y_t}{Y^*_t}\right)^{\frac{1}{\gamma}}.$$

On the other hand, we showed above that for all $u \in [0, \bar{u}]$

$$\mathcal{P}_{\ell,t}(u) \;=\; \mathcal{P}^*_{\ell,t}(u) \left(\frac{Y_t}{Y^*_t}\right)^{\frac{1}{\gamma}}.$$

Combining these two inequalities yields

$$\frac{\mathcal{P}_{h,t}(\bar{u})}{\mathcal{P}^*_{h,t}(\bar{u})} \;>\; \frac{\mathcal{P}_{\ell,t}(\bar{u})}{\mathcal{P}^*_{\ell,t}(\bar{u})}.$$

Moreover, the wage for workers of age $\bar{u}_p$ is

$$W_{h,t}(\bar{u}_p) \;=\; \mathcal{P}_{h,t}(\bar{u})\, a(\bar{u}, \bar{u}_p),$$

which implies

$$\frac{W_{h,t}(\bar{u}_p)}{W^*_{h,t}(\bar{u}_p)} \;>\; \frac{W_{\ell,t}(\bar{u}_p)}{W^*_{\ell,t}(\bar{u}_p)} \;\Leftrightarrow\; \frac{W_{h,t}(\bar{u}_p)}{W_{\ell,t}(\bar{u}_p)} \;>\; \frac{W^*_{h,t}(\bar{u}_p)}{W^*_{\ell,t}(\bar{u}_p)}.$$

This shows that the skill premium rises for workers of age $\bar{u}_p$ early on during the transition.

We now show the increase is even more pronounced for younger workers. For any $t \in (t_0, \min\{\bar{t}, t_0 + \underline{u}\})$, the envelope theorem applied to $\ln W_{h,t}(u_p)$ implies

$$\frac{W'_{h,t}(u_p)}{W_{h,t}(u_p)} \;=\; \frac{\dfrac{\partial a\big(\mathcal{U}_{h,t}(u_p), u_p\big)}{\partial u_p}}{a\big(\mathcal{U}_{h,t}(u_p), u_p\big)} \;+\; \beta_h \;<\; \frac{\dfrac{\partial a\big(\mathcal{U}^*_{h,t}(u_p), u_p\big)}{\partial u_p}}{a\big(\mathcal{U}^*_{h,t}(u_p), u_p\big)} \;+\; \beta_h \;=\; \frac{{W^*}'_{h,t}(u_p)}{W^*_{h,t}(u_p)}.$$

The middle inequality relies on the fact that skilled workers shift toward newer technologies (a consequence of ii above), and young workers have a comparative advantage in these technologies.

Likewise, for any $t \in (t_0, \min\{\bar{t}, t_0 + \underline{u}\})$, the envelope theorem applied to $\ln W_{\ell,t}(u_p)$ implies

$$\frac{W'_{\ell,t}(u_p)}{W_{\ell,t}(u_p)} \;=\; \frac{\dfrac{\partial a\big(\boldsymbol{\mathcal{U}}_{\ell,t}(u_p), u_p\big)}{\partial u_p}}{a\big(\boldsymbol{\mathcal{U}}_{\ell,t}(u_p), u_p\big)} \;+\; \beta_\ell \;=\; \frac{\dfrac{\partial a\big(\boldsymbol{\mathcal{U}}^*_{\ell,t}(u_p), u_p\big)}{\partial u_p}}{a\big(\boldsymbol{\mathcal{U}}^*_{\ell,t}(u_p), u_p\big)} \;+\; \beta_\ell \;=\; \frac{W^{*\prime}_{\ell,t}(u_p)}{W^*_{\ell,t}(u_p)}.$$

This shows that, for any $t \in (t_0, \min\{\bar{t}, t_0 + \underline{u}\})$, the change in the college premium by worker age, relative to its BGP level

$$\Delta(u_p) \;=\; \ln\frac{W_{h,t}(u_p)}{W_{\ell,t}(u_p)} - \ln\frac{W^*_{h,t}(u_p)}{W^*_{\ell,t}(u_p)}$$

is more pronounced among younger workers, since

$$\Delta'(u_p) \;=\; \underbrace{\left(\frac{W'_{h,t}(u_p)}{W_{h,t}(u_p)} - \frac{W^{*\prime}_{h,t}(u_p)}{W^*_{h,t}(u_p)}\right)}_{<0} \;-\; \underbrace{\left(\frac{W'_{\ell,t}(u_p)}{W_{\ell,t}(u_p)} - \frac{W^{*\prime}_{\ell,t}(u_p)}{W^*_{\ell,t}(u_p)}\right)}_{=0} < 0,$$

and this is positive for all workers since $\Delta(\bar{u}_p) > 0$. □