Technological Transitions with Skill Heterogeneity Across Generations*

Rodrigo Adão†
University of Chicago
Booth School of Business

Martin Beraja‡
Massachusetts Institute of Technology

Nitya Pandalai-Nayar§
University of Texas at Austin

July 13, 2019

Abstract

How do economies adjust to technological innovations? We develop a theory where overlapping generations of workers are heterogeneous over a continuum of technology-specific skills. A worker skill-type results from a costly investment upon entry. Given a type's technology-specific wage, they self-select into a technology. We show that this economy can be represented as a q-theory of skill investment. This allow us to sharply characterize the transitional dynamics and welfare implications of a technology-improving innovation. Economies where technology-skill specificity is higher have a slower adjustment following the innovation because the larger increases in relative wages induce larger changes in the skill-distribution across generations. We use our theory to study how German regions adjusted to the introduction of broadband internet in the 2000s. In early adopting regions, employment in cognitive-intensive occupations increased for young workers but not for old workers. These results suggest that skill-specificity is high and that the supply of skills is more elastic at longer horizons. Ignoring the adjustment across generations by focusing on short- or long-run changes alone biases the average and distributional welfare implications of technological innovations. Such biases are severe precisely in economies with high skill-specificity and large observed increases in inequality.

*We are grateful to Masao Fukui and Michelle Lam for excellent research assistance. For their valuable suggestions, we thank Daron Acemoglu, Arnaud Costinot, Ariel Burstein, V.V. Chari, Oded Galor, Pablo Kurlat, Frank Levy, Fabrizio Perri, Venky Venkateswaran, and Fabrizio Zilibotti. We also thank seminar participants at Duke, MIT, NYU, Notre-Dame, UCSD, UT Austin, WashU, St. Louis Fed, UCLA, USC Marshall, IIES Stockholm, the Barcelona GSE Summer Forum, and the NBER SI Income Distribution and Macroeconomics group for helpful comments. Data access was provided by the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB). We thank Peter Brown and Clare Dingwell of the Harvard Economics Department, for facilitating access to IAB data. All results based on IAB microdata have been cleared for disclosure to protect confidentiality.

†Email: rodrigo.adao@chicagobooth.edu. Web: sites.google.com/site/rradao

‡Email: maberaja@mit.edu. Web: economics.mit.edu/faculty/maberaja

§Email: npnayar@utexas.edu. Web: www.nityapandalainayar.com

1 Introduction

New technologies are the key drivers of increases in living standards over long horizons. Yet, more recently, a literature has shown that they may have strong distributional consequences at shorter horizons.[1] This dichotomy is particularly relevant when the adjustment happens over generations. In such cases, one risks missing the average and distributional welfare implications of new technologies when only considering their long-run or short-run consequences. In this paper, we develop a new theory to study technological transitions that emphasizes skill-heterogeneity within and across generations. We then use it to empirically assess the implications of recent skill-biased innovations.

The theory has three distinct features. First, there are overlapping generations of workers with stochastic lifetimes, as in Yaari (1965) and Blanchard (1985). Second, within each generation, workers are heterogeneous over a continuum of skill types. A type determines the worker's productivity in the two technologies of the economy, as in Roy (1951). Given their type's technology-specific wages, workers self-select into one of the two technologies at each point in time. The output of the two technologies is then combined to produce a final consumption good. Third, given the expected future path of wages, workers make a costly investment upon entering the labor market that determines their skill type, similar to Chari and Hopenhayn (1991) and Caselli (1999). This gives rise to differences in technology-specific skill heterogeneity across generations.

The equilibrium of this economy is a joint path for the skill distribution, the assignment of skill types to technologies, and relative technology-specific wages. Our first result establishes that the approximate equilibrium dynamics of this economy can be represented as a q-theory of skill investment, where q is the present-discounted value of log-relative wages and the skill distribution plays the role of the pre-determined variable.[2] This then determines the assignment of workers and relative wages at each point in time.

Our second result derives in closed-form the transitional dynamics following a one-time, permanent increase in the productivity of all skill types employed in one of the technologies. We refer to this as a skill-biased technological innovation. The logic of the economy's adjustment follows immediately from the q-theory representation of the equilibrium. The relative productivity increase leads to an increase in the demand for workers and relative wages in the improved technology. On impact, marginal skill-types now sort into that technology. The extent to which they do so crucially depends on how different skill types are in their relative technology-specific productivity. In our theory, this is controlled by a parameter of technology-skill specificity which determines the short-run skill supply elasticity (η). The increase in future relative wages leads younger entering generations to invest in skills that

[1] See Durlauf and Aghion (2005) for a review of the literature on the impact of new technologies and innovation on long-run living standards. See Acemoglu and Autor (2011) and Autor and Salomons (2017) for a review of the literature documenting the impact of new technologies on employment and wages of workers associated with different skills, occupations, industries, and firms.

[2] See Tobin (1969) and Hayashi (1982) for the original q-theory of capital investment.

are complementary to the improved technology. The extent to which they do so depends on a parameter governing the cost of investing in technology-specific skills which determines the skill supply elasticity at longer horizons (ψ). Along the transition, q falls and relative output increases as younger generations replace older generations, expanding the supply of skills that became more valuable after the shock.

Our third result establishes that an economy where technology-skill specificity is higher has a slower adjustment path to the new long-run equilibrium. Specifically, it has more persistent dynamics of q and relative output as measured by their cumulative impulse responses. The q-theory analogy again delivers the intuition for this result. When technology-skill specificity is higher (and η is lower), following the technological innovation, the reallocation of worker skill-types across technologies is smaller and, therefore, the increase in relative wages is larger. This implies that younger entering generations have stronger incentives to invest in the skills that became more valuable. As a result, there are larger differences in skill heterogeneity across generations. Thus, the economy's adjustment is slower because larger changes in the skill distribution take place as younger generations replace older generations. Importantly, a higher long-run skill supply elasticity (ψ) amplifies the impact of technology-skill specificity on the persistence of the economy's adjustment.

We conclude the theoretical part of the paper by analyzing how the persistence of the adjustment affects the average and distributional welfare consequences of a skill-biased technological innovation. In our first thought experiment, we ignore the adjustment across generations by setting the long-run skill supply elasticity to zero (i.e., $\psi = 0$). Since this parameter does not affect responses on impact, we obtain the same short-run responses in all outcomes. This is equivalent to analyzing changes triggered by a skill-biased innovation in a static assignment model. As such, it misses the inequality decline and the output increase that happens at longer horizons. This implies that it understates the average welfare benefits and overstates the lifetime welfare inequality increase caused by the innovation. Crucially, the magnitude of these biases are larger in economies that exhibit larger short-run increases in inequality due to higher technology-skill specificity. In our second thought experiment, we set η to match a given relative output increase in the long-run while ignoring the adjustment across generations (i.e., $\psi = 0$). Again, because of our third result, this long-run calculation ignores the slow adjustment of skills across generations and thus overstates the average welfare benefits and understates the lifetime welfare inequality increase.

In the second part of the paper, we empirically assess the role that our two main theoretical mechanisms play in the adjustment of economies to recent skill-biased technological innovations. We start by connecting the parameters governing technology-skill specificity and skill investment cost to observable dynamic responses of worker allocations within and between generations. We do so by exploiting the closed-form expressions for the economy's transitional dynamics. Intuitively, for older generations of workers with a given skill distribution, the innovation-induced employment reallocation is larger if skills are less specific to each technology (i.e., η is higher). Relative to older generations, younger workers adjust their skills in response to the technological innovation. This generates between-generation differences in relative employment that are larger whenever the cost of investing in skills is lower (i.e. ψ is higher).

We apply these insights to study the experience of Germany between 1997 and 2014. We first document that the use of two important new technologies, computers and internet, is heavily biased towards cognitive-intensive occupations. Conditional on a worker's occupation, the use of these technologies does not vary across generations. We rely on these observations to define cognitive-intensive occupations as the set of production activities disproportionately augmented by innovations in these two technologies. We then estimate how the cognitive intensity of an occupation affects its employment growth for different generations of workers. We find that overall employment growth was statistically significantly higher in more cognitive intensive occupations. However, we obtain very different estimates across worker generations. While the differential employment growth in more cognitive intensive occupations was close to zero for older generations, it was positive for younger generations.

This evidence speaks to the differential adjustment across generations to skill-biased innovations, but it is silent about the full transitional dynamics implied by our theory. For this reason, we next analyze the dynamic response to one particular cognitive-biased technological innovation: the introduction of broadband internet in the early 2000s. There are three reasons to focus on this particular innovation in Germany. First, the adoption of broadband internet was fast: the share of households using it increased from 0% in 2000 to over 90% in 2009. Second, it was spatially heterogeneous: across German districts in 2005, the mean share of households with broadband internet access was 76% and the standard deviation was 16%. Third, following Falck, Gold, and Heblich (2014), it is possible to isolate exogenous spatial variation in adoption timing coming from the suitability of pre-existing local telephone networks for broadband internet transmission. Taken together, these reasons imply that the introduction of broadband internet resembles a one-time permanent shock like the one studied in our theory, thus allowing the estimation of impulse response functions.

We estimate the impact of higher broadband internet penetration in 2005 on labor market outcomes across German districts at different horizons. As suggested by the theory, we focus on the differential employment and payroll growth in more cognitive intensive occupations for different worker generations. Compared to late adopting regions, we find that in early adopting regions relative employment and payroll in more cognitive intensive occupations was constant before 2005 but slowly increased afterwards. However, the estimates are again different for older and young generations. The impact on relative employment is small and nonsignificant for older generations at all horizons, but it is positive and statistically significant for younger generations after 2005.

Our empirical results suggest that skill supply elasticity is close to zero at short horizons but it is positive at longer horizons. Through the lens of our theory, such patterns arise whenever skills are very specific to cognitive intensive activities (i.e., η close to zero) and the cost of accumulating skills is smaller for younger entering generations (i.e., ψ is positive). As discussed above, this is precisely the environment that is likely to lead to substantial welfare biases from ignoring the slow adjustment across generations.

To quantify this, we calibrate our model to match our estimated dynamic responses for Germany. We consider a skill-biased innovation that increases the employment share of a technology from 20 percent to 40 percent across long-run equilibria. We find that the consumption equivalent average welfare increase across all generations is 46 percent and the lifetime welfare inequality increase is 39 percent. We then compute the biases of ignoring the adjustment across generations implied by our theory. If we match the same short-run increase in the employment share and only consider the changes on impact, we instead find a lower average welfare increase of 31 percent and a larger inequality increase of 76 percent. If we match the same long-run increase in relative employment and only consider changes across steady states, we instead find a higher average welfare increase of 55 percent and a lower inequality increase of 30 percent. These biases are smaller in economies with lower skill-specificity because the economy's adjustment is faster.

Related Literature. Our paper is related to several strands of the literature. A long literature has analyzed the labor market consequences of technological innovations. We depart from the canonical framework in Katz and Murphy (1992) by modeling the supply of skills across technologies at different time horizons. Specifically, given the skill distribution at any point in time, the short-run skill supply to each technology arises from the static sorting decision of workers. This static assignment structure has been used in a recent literature analyzing how labor markets respond to a variety of shocks – e.g, Costinot and Vogel (2010), Acemoglu and Autor (2011), Hsieh, Hurst, Jones, and Klenow (2013), Burstein, Morales, and Vogel (2016), and Adão (2016). In addition, our theory entails slow-moving changes in skill supply that arise from the entry of young generations with different skills than those of previous generations, as in Chari and Hopenhayn (1991), Caselli (1999) and Galor and Moav (2002).[3] We show that the combination of these features yields tractable expressions for the equilibrium dynamics that resemble a q-theory of skill investment. We exploit the parsimony of our theory to establish that higher levels of skill-technology specificity and long-run skill supply elasticity generate slower adjustments following skill-biased innovations, affecting its average and distributional welfare consequences at different horizons. We then link the two margins of skill supply in our theory to observable dynamic responses of labor market

[3]Recent papers have documented that demand-driven shocks in relative wages affect educational attainment decisions – e.g., Atkin (2016) and Charles, Hurst, and Notowidigdo (Forthcoming). Our theory builds on the insight in Ben-Porath (1967) that workers make the bulk of their skill investment early in the life cycle, implying that young workers have a higher elasticity to changes in relative wages than old workers (e.g., Lee and Wolpin (2006)). We make the extreme assumption that the skills of old workers are fixed and, therefore, abstract from dynamic responses stemming from re-training. At the cost of some realism, the model gains substantially in terms of its tractability, allowing us to analytically characterize the evolution of the skill distribution in response to technological shocks.

outcomes within and across generations. Our empirical application indicates that separately allowing for these two forces is important in the context of Germany.

The main source of dynamics in our theory is the endogenous change in the supply of technology-specific skills over time. Several papers have proposed alternative sources of dynamics to study the transition following technological innovations, including sluggish labor mobility across sectors (Matsuyama, 1992), technology diffusion across firms (Atkeson and Kehoe, 2007), firm-level investment in R&D (Atkeson, Burstein, and Chatzikonstantinou, 2018), and endogenous creation of new tasks for labor in production (Acemoglu and Restrepo, 2018) or permanent changes in the returns to wealth accumulation following increases in automation (Moll, Rachel, and Restrepo, 2019). Our paper complements this literature by analyzing empirically and theoretically how the endogenous dynamics of skill heterogeneity across generations shape the economy's adjustment to skill-biased technological innovations.

An extensive literature has estimated the distributional consequences of new technologies – for a review, see Acemoglu and Autor (2011). Our empirical analysis follows the literature showing how new technologies affect occupations with different task intensity – e.g., Autor, Levy, and Murnane (2003), Autor and Dorn (2013) and Acemoglu and Restrepo (2017). As in Akerman, Gaarder, and Mogstad (2015), we exploit regional characteristics to estimate the labor market consequences of broadband internet adoption. While they focus on the impact of this technology on the educational composition of employment in Norwegian firms, we estimate its effect on the occupation composition of employment in German regional labor markets. Similar to Card and Lemieux (2001) and Autor and Dorn (2009), we find that the impact of new technologies varies across worker generations. We complement this literature by showing that such short-run empirical analysis alone may miss part of the impact of new technologies on average welfare and lifetime inequality in economies with high skill-specificity, but it is an essential input in the measurement of the two main determinants in our model of the transitional dynamics and welfare changes triggered by the arrival of new technologies.

Our paper is also related to the literature analyzing structural transformation in the form of long-run worker reallocation across sectors – e.g., Ngai and Pissarides (2007), Buera and Kaboski (2012), Herrendorf, Herrington, and Valentinyi (2015) and, for a review, Herrendorf, Rogerson, and Valentinyi (2014). Recently, Young (2014) and Lagakos and Waugh (2013) show that endogenous skill-sector sorting affects the process of structural transformation. Moreover, a number of papers have also emphasized the adjustment across generations. Kim and Topel (1995) document that the expansion of the manufacturing sector in Korea was driven almost entirely by new, young entrants to the labor force. Porzio and Santangelo (2019) documents substantial variation across countries in the extent to which the reallocation out of agriculture happens within- or between-cohorts. Hobijn, Schoellman, and Vindas (2019) also document large between-generation differences in worker reallocation across countries, extending the analysis to agriculture, manufacturing and services. Relative to this literature, we make three contributions. First, we provide a tractable theory to analyze the role of skill heterogeneity within and across generations in the transitional dynamics following to technological innovations. Second, we estimate impulse response functions to a well-identified technological innovation in Germany and show how they discipline the key parameters of our theory. Third, we point out which features of the economy (e.g., skill-specificity) lead to slow adjustment dynamics and, as result, large biases from welfare calculations that ignore them.

Outline. Our paper is organized as follows. Section 2 presents our model and establishes the q-theory representation of its equilibrium. In Section 3, we analyze the dynamic adjustment to skill-biased technological innovation. Section 4 links our main theoretical mechanisms to responses in observable outcomes for different generations of workers. Section 5 presents our empirical analysis of the effect of skill-biased innovations in Germany. Section 6 presents our quantitative analysis. Section 7 concludes.

2 A Model of Skilled-biased Technological Transitions

We start by proposing a model with overlapping generations of workers that make forward-looking decisions to invest on a continuum of skills upon entry. Conditional on their skills, workers self-select to work with different technologies throughout their lives. Our model yields a q-theory of skill investment that determines the equilibrium path of the wage and output of workers employed with different technologies. As a function of these outcomes, we characterize the evolution of the economy's skill distribution and skill-technology assignment.

2.1 Environment

We consider a closed economy in continuous time indexed by t. There is a single final output whose production uses the input of two intermediate goods. The production technology of each intermediate good requires workers to perform a technology-specific task bundle. We denote the two technologies as high-tech ($k = H$) and low-tech ($k = L$). There is a continuum of worker skill types, $i \in [0, 1]$, that determine the worker's productivity with each production technology.

Final Product. Production of the final product is a CES aggregator of the two intermediate inputs:

$$Y_t = \left[(A_t Y_{Ht})^{\frac{\theta-1}{\theta}} + (Y_{Lt})^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} \tag{1}$$

where $\theta > 0$ is the demand elasticity of substitution between the low-tech and the high-tech intermediate products, and A_t is a shifter of the relative productivity of the high-tech input.

Conditional on the price of intermediate inputs, the cost minimization problem of firms producing final product implies that the relative spending on the high-tech input is

$$y_t = \left(\frac{\omega_t}{A_t}\right)^{1-\theta}, \tag{2}$$

where $\omega_t \equiv \omega_{Ht}/\omega_{Lt}$ is the relative price of the high-tech input. We normalize the price of the low-tech input to one, $\omega_{Lt} \equiv 1$.[4]

We consider a competitive environment, so that profit maximization implies that the equilibrium price of the final output is

$$P_t = (1 + y_t)^{\frac{1}{1-\theta}}. \tag{3}$$

In the model, y_t measures the relative importance of high-tech production in total output, and ω_t is the relative value of high-tech production. In the next section, we analyze how these outcomes adjust following unanticipated permanent changes in A_t. As in Katz and Murphy (1992), A_t is a shifter of the relative productivity of the high-tech input. We now describe the structure of skill heterogeneity in our model.

Allocation of skills to technology. We assume that a worker's skill type determines her productivity with the two technologies in the economy. For a worker of type i, $\alpha(i)$ is the overall productivity and $\sigma(i)$ is the differential productivity in high-tech production. Specifically, we assume that the production function of the low-tech good is

$$X_{Lt} = \int_0^1 \alpha(i) s_{Lt}(i) di, \tag{4}$$

and that of the high-tech good is

$$X_{Ht} = \int_0^1 \alpha(i)\sigma(i) s_{Ht}(i) di, \tag{5}$$

where $s_{kt}(i)$ is the density function of workers employed with technology k at time t.

We assume a competitive labor market with zero profit in low-tech and high-tech production. In equilibrium, the wage rates of skill type i with the H and L technologies are respectively given by

$$w_{Ht}(i) = \omega_t \sigma(i) \alpha(i) \quad \text{and} \quad w_{Lt}(i) = \alpha(i). \tag{6}$$

As in Roy (1951), workers self-select across technologies to maximize labor income. Thus,

[4]In Appendix C.2, we show that the relative demand in (2) can be derived from alternative microfoundations. Specifically, it arises in a small open economy in a gravity trade model, as in Eaton and Kortum (2002) and Arkolakis, Costinot, and Rodríguez-Clare (2012), and in an environment with a continuum of heterogeneous firms choosing technology adoption.

the wage of a worker with skill type i is

$$w_t(i) = \max\{\omega_t \sigma(i), 1\}\alpha(i). \tag{7}$$

Recent papers have considered similar structure of endogenous sorting of workers to different technologies – e.g., Acemoglu and Autor (2011), Costinot and Vogel (2010), Adão (2016). As each point in time t, the sorting decision in (7) generates a mapping of skill types to technology. Such a mapping depends on the relative value of one unit of effective labor units employed in high-tech production, ω_t. In the model, ω_t is a natural measure of inequality as it captures the relative wage rate of skill type's employed in different technologies conditional on their productivity. The sorting decision also depends the exogenous relative productivity of different skill types in the H technology. That is, the function $\sigma(i)$ that measures type i's comparative advantage in high-tech production. Without loss of generality, it is always possible to order skill types such that high-i types have a higher relative productivity in high-tech production. Thus, the slope of the function $\sigma(i)$ captures the degree skill-technology specificity: how much the relativity productivity in technology H varies across skill types. To simplify the analysis, we assume that $\sigma(i)$ is strictly increasing.

Skill investment. We consider an overlapping generations setting in which the birth and death of workers follows a Poisson process with rate δ. At each point in time, workers use their labor earnings to purchase the final good. Utility is the present value of a logarithmic flow utility discounted by a rate ρ. For a worker of type i born at time t, lifetime utility is

$$V_t(i) = \int_t^{\infty} e^{-(\rho+\delta)(s-t)} log(w_s(i))ds. \tag{8}$$

Crucially, as in Chari and Hopenhayn (1991) and Caselli (1999), we allow workers to acquire different skills at birth taking into account the value of future earnings streams. Given the wage function $\{w_s(i)\}_{s>t}$, workers born at time t can pay a utility cost to select a lottery $\tilde{s}_t(i)$ over skill types. If they do not pay the cost, their type is drawn from an exogneous distribution of innate ability, $\bar{s}_t(i)$.[5] A worker's type is then fixed during their lifetime. Intuitively, workers can pay an ex-ante cost to invest on human capital that allows them to target specific skill types. The skill investment may occur through schooling or on-the-job while workers are young. Yet, we assume that there is uncertainty regarding the return of the investment in terms of the specific skill realization.[6]

Formally, we assume that the cost of the lottery is proportional to the Kullback-Leibler

[5] Appendix C.1 considers the case of "learning from others" by allowing $\bar{s}_t(i)$ to depend on the current skill distribution in the economy $s_t(i)$. As in Chari and Hopenhayn (1991), this introduces complementarity in the skill investment decision of different generations.

[6] Our approach is consistent with the evidence in Carneiro, Heckman, and Vytlacil (2011) that, conditional on observable educational investment, individuals with different unobserved skills have heterogeneous educational returns. Our model treats this heterogeneity in unobserved skills through the uncertainty of the type realization. This approach is isomorphic to an environment where heterogeneous individuals in terms of skill cost acquisition select their skill type.

divergence between the lottery and the baseline distribution $\bar{s}_t(i)$.[7] Specifically, the cost of lottery $\tilde{s}_t(i)$ is

$$\frac{1}{\psi}\int_0^1 log\left(\tilde{s}_t(i)/\bar{s}_t(i)\right)\tilde{s}_t(i)di, \qquad \psi > 0.$$

Thus, the skill investment problem is

$$\max_{\tilde{s}_t(.):\int_0^1 \tilde{s}_t(i)di=1} \int_0^1 \left(V_t(i) - \frac{1}{\psi} log\left(\frac{\tilde{s}_t(i)}{\bar{s}_t(i)}\right)\right)\tilde{s}_t(i)di. \tag{9}$$

The parameter ψ governs the cost of targeting particular skill types. In the limit when $\psi \to 0$, the cost of targeting a particular skill is infinite and the economy's skill distribution does not respond to changes in the lifetime earnings of different skill types. Whenever $\psi > 0$, the optimal lottery $\tilde{s}_t(i)$ endogenously responds to the relative present discounted value of different skill types, $V_t(i)$.

Finally, for our main analysis, we will only consider a fix reference distribution $\bar{s}(i)$. In an extension section we also consider how the analysis changes when we allow for a form of learning-from-others such that $\bar{s}_t(i) = (s_t(i))^{\gamma}(\bar{s}(i))^{1-\gamma}$. Here γ governs the strength of this learning-from-others effect: when γ is higher, it becomes easier for workers to target skills that are already abundant in the economy.

Equilibrium. The assumption that only new cohorts have the ability to choose a skill lottery implies that the evolution of the skill distribution $s_t(i)$ follows the Kolmogorov-Forward equation,

$$\frac{\partial s_t(i)}{\partial t} = -\delta s_t(i) + \delta \tilde{s}_t(i). \tag{10}$$

Finally, the economy's equilibrium must satisfy market clearing for all t. By Walras law, it is sufficient to only consider the relative demand and supply of the high-tech intermediate input:

$$y_t = \omega_t x_t \tag{11}$$

where y_t is given by (2) and x_t is the ratio of the production functions in (4)–(5).

Definition 1 (Competitive Equilibrium) *Given an initial skill distribution $s_0(i)$, a competitive equilibrium is a path of the worker-technology assignment $\{G_t(i) : i \in [0,1] \to \{H, L\}\}_t$, the skill distribution $\{s_t(i)\}_t$, the skill lottery $\{\tilde{s}_t(i)\}_t$, the relative price and ideal price index $\{\omega_t, P_t\}_t$, such that*

[7] This cost function has a long tradition in macroeconomics that uses it as a tractable form of comparing distance between distributions – e.g., frameworks with rational inattention (Sims (2003)) and model uncertainty (Hansen and Sargent (2008)). As discussed below, in our setting, this function yields the continuous type analog of the skill acquisition solution in an environment in which worker's ability to acquire a discrete number of skills follows a Type 1 extreme-value distribution.

1. *Given ω_t, the worker-technology assignment is determined by the worker self-selection decision in (7).*

2. *The skill distribution $s_t(i)$ satisfies the Kolmogorov-Forward equation (10) with the skill lottery $\tilde{s}_t(i)$ given by the skill investment decision of new cohorts implied by (9).*

3. *The markets for high-tech and low-tech goods clear by satisfying (11). The final good market clears by satisfying (3).*

2.2 Static and Dynamic Equilibrium Conditions

We now proceed to characterize the equilibrium in two steps. First, we consider the static conditions that, given the skill distribution $s_t(i)$ at time t, determines the relative wage and output of the high-tech input, $\{\omega_t, y_t\}_t$, that clears the intermediate good market. Second, we consider dynamic conditions that, given the path of the relative wage $\{\omega_t\}_t$, determine the optimal skill lottery chosen by young cohorts $\{\tilde{s}_t(i)\}_t$ and the evolution of the skill distribution $\{\tilde{s}_t(i)\}_t$.

Static Equilibrium Conditions. The endogenous sorting decision in (7) determines the assignment of skill types to technologies. We graphically represent this decision in Figure 1. The solid and dashed lines represent the potential log-wage of workers of type $i \in [0,1]$ if employed, respectively, in high- and low-tech production. Equation (7) implies that types self-select to work with the technology that yields the highest labor earnings. Thus, high-i (low-i) types receive higher relative earnings in high-tech (low-tech) production and choose to be employed with that technology. In equilibrium, the assignment is described by a threshold l_t characterizing the type that is indifferent between working with any of the two technologies. [8] That is,

$$1 = \omega_t \sigma(l_t). \tag{12}$$

We summarize this discussion in the following lemma.

Lemma 1 (Equilibrium Assignment) *Worker types $i \leq l_t$ are employed in low-tech production with labor income of $w_t(i) = \alpha(i)$. Worker types $i > l_t$ are employed in high-tech production with labor income of $w_t(i) = \omega_t \sigma(i) \alpha(i)$. The threshold is determined by the indifference condition in (12).*

Lemma 1 links the relative wage ω_t to the allocation of skill types across technologies. Condition (12) is central to understand how technological shocks affect the allocation of workers across technologies. It implies that the inverse elasticity of $\sigma(i)$ determines how much workers reallocate across technologies in response to changes in the relative wage. We

[8] Notice that, since $\sigma(i)$ is strictly increasing, the two curves in Figure 1 must cross at most once. Since the demand function in (2) goes to infinity as ω_t goes to zero, there is positive employment with both technologies in equilibrium and, therefore, the two curves in Figure 1 must cross at least once. Figure 1 illustrates the case in which high-i types are more productive in absolute terms, receiving higher earnings if employed with any of the two technologies. Adão (2016) shows that, depending on the shape of $\alpha(i)$, there are different possible configurations for Figure 1 which determine the correlation between income and technology allocation across skill types.

refer to the elasticity of $\sigma(i)$ as the technology-skill specificity. The inverse elasticity of $\sigma(i)$ plays the role of short-run skill supply across technologies, which we formally define as

$$\eta \equiv \left|\frac{\partial \log l_t(\omega_t)}{\partial \log \omega_t}\right| = \left(\frac{\partial \log \sigma(l_t)}{\partial \log i}\right)^{-1}.$$

where $l_t(\omega_t)$ is the implicit function defined by (12).

The skill-technology assignment in Lemma 1 determines the relative supply of high-tech production as a function of the relative wage ω_t. Conditional on the skill distribution $s_t(i)$, equations (4)–(5) imply that the relative supply is

$$x_t(\omega_t, s_t) = \frac{\int_{l_t(\omega_t)}^{1} \sigma(i)\alpha(i)s_t(i)di}{\int_0^{l_t(\omega_t)} \alpha(i)s_t(i)di}. \tag{13}$$

The relative price ω_t is then determined by the market clearing condition in (11). We illustrate this condition in Panel B of Figure 1. Whenever ω_t is low, relative demand for input H is high, but its relative supply is low due to the small share of types employed in high-tech production. Whenever ω_t is high, the opposite is true. In equilibrium, relative demand and supply are equalized. The following lemma formalizes the existence of a unique equilibrium value of ω_t for any given distribution $s_t(i)$.

Lemma 2 (Equilibrium Threshold) *Given the skill-type distribution $s_t(i)$ at time t, there is a unique equilibrium relative wage such that*

$$A_t^{\theta-1}\omega_t^{-\theta}\int_0^{l_t(\omega_t)} \alpha(i)s_t(i)di = \int_{l_t(\omega_t)}^{1} \alpha(i)\sigma(i)s_t(i)di, \tag{14}$$

and $l_t(\omega_t)$ is the implicit function defined by (12).

Proof. See Appendix A.1. ■

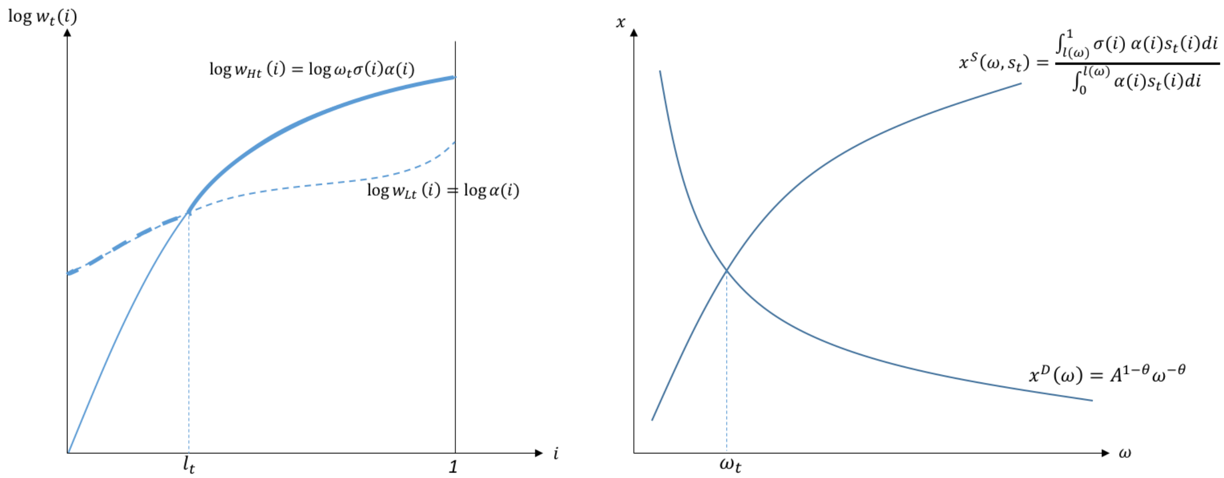
Dynamic Equilibrium Conditions. We now turn to the entrant's forward-looking problem of choosing their skill lottery $\tilde{s}_t(i)$, conditional on the future path of the relative technology-specific wage $\{\omega_s\}_{s>t}$. The solution of the maximization problem in (9) immediately yields the following optimal lottery.

Lemma 3 (Equilibrium Lottery) *Define $log\left(Q_t(i)\right) \equiv \int_t^\infty e^{-(\rho+\delta)(s-t)} \max\{log\left(\omega_s\sigma(i)\right), 0\}ds$. The equilibrium lottery satisfies:*

$$\tilde{s}_t(i) = \frac{\bar{s}_t(i)\alpha(i)^{\frac{\psi}{\rho+\delta}}Q_t(i)^{\psi}}{\int_0^1 \bar{s}_t(j)\alpha(j)^{\frac{\psi}{\rho+\delta}}Q_t(j)^{\psi}dj} \tag{15}$$

Proof. See Appendix A.2. ■

Figure 1: Static Equilibrium Conditions



Panel A: Equilibrium Assignment

Panel B: Market Clearing

The optimal lottery in (15) is an extension of the multinomial logit function for a continuum of types. It shows that the investment on high-i types is a function of the present value of the relative wage in high-tech production as captured by Q_t. The parameter ψ governs the sensitivity of the optimal lottery to changes in relative lifetime earnings. To see this more clearly, consider the stationary equilibrium with $\omega_t = \omega$ such that

$$s(i) = \tilde{s}(i) = \frac{\bar{s}(i)W(i)^{\psi}}{\int_0^1 \bar{s}(j)W(j)^{\psi} dj} \tag{16}$$

where $W(i) = (\alpha(i) \max\{\omega\sigma(i), 1\})^{\frac{1}{\delta+\rho}}$ is the present value of the utility flow of skill type i.

In this case, the skill distribution is a constant-elasticity function of relative wages across types, where the elasticity is $\frac{\psi}{\rho+\delta}$.[9] Thus, a higher ψ implies that the long-run supply of high-i types is more sensitive to changes in the relative wage in high-tech production. Accordingly, we refer to ψ as the long-run skill supply across technologies, which we formally define as

$$\psi \equiv \frac{\partial \log s(i)/s(i')}{\partial \log W(i)/W(i')}.$$

[9]Notice that the log-run equilibrium of our model is a generalization with a continuum of types of the extension of the assignment model in Acemoglu and Autor (2011) with endogenous skill supply – see Section 4.6 in Acemoglu and Autor (2011). In our framework however, along the transitional equilibrium, the skill-distribution differs from this stationary distribution.

2.3 Equilirium Characterization: Q-Theory of Skill Investment

We now combine the static equilibrium condition in (14) and dynamic equilibrium condition in (15) to characterize the equilibrium path of the relative technology-specific wage rates, $\{\omega_t\}_t$ and the skill distribution $\{s_t(i)\}_t$. We begin by noting that the only endogenous state variable in our model is the skill distribution $s_t(i)$. In addition, the only control variable is the optimal lottery, $\tilde{s}_t(i)$, that is a function of the path of relative wage $\{\omega_s\}_{t>s}$. This implies that the dynamic equilibrium of the economy entails a fixed-point problem, since the market clearing condition in (14) implies that ω_t is a function of $s_t(i)$ through its effect on the relative high-tech output supply.

Our first result approximates the equilibrium solution of this fixed-point problem by considering a log-linear expansion around the stationary equilibrium. Proposition 1 establishes that one does not need to keep track of the whole skill-distribution in order to solve for the log-linearized equilibrium, which simplifies the fixed-point problem considerably. Instead, denoting with " ˆ " variables in log-deviations from the stationary equilibrium, there is a system of differential equations that characterizes the joint dynamics of $\{\hat{q}_t, \hat{y}_t\}$, given an initial $\hat{x}_0$, where $log(q_t) \equiv \int_t^\infty e^{-(\rho+\delta)(s-t)} log(\omega_s) ds$ is the present discounted value of the log-relative wage.

The resulting system of differential equations is a rather standard one in macroeconomics. Relative output, $\hat{y}_t$, acts as a state variable whose law of motion needs to be solved backward. The optimal lottery is the control variable that needs to be solved forward. The present value of future relative technology-specific wage, q_t, is the only endogenous variable affecting the optimal lottery choice, so it is the only variable that needs to be solved forward.[10]

Proposition 1 (Q-theory of Skill Investment)

1. *Consider the initial condition $\hat{y}_0$ and the terminal condition $\lim_{t\to\infty} \hat{y}_t = 0$. The dynamic system describing the joint evolution of $\{\hat{q}_t, \hat{y}_t\}$ is*

$$\frac{\partial \hat{y}_t}{\partial t} = -\delta \hat{y}_t + \frac{\theta - 1}{\theta + \kappa\eta} \delta \psi \hat{q}_t \tag{17}$$

$$\frac{\partial \hat{q}_t}{\partial t} = (\rho + \delta)\hat{q}_t + \frac{1}{\theta - 1} \hat{y}_t, \tag{18}$$

   where κ is a positive constant.

2. *The equilibrium for $\{\hat{q}_t, \hat{y}_t\}$ is saddle-path stable and given by:*

$$\hat{y}_t = \hat{y}_0 e^{-\lambda t} \quad \textit{and} \quad \hat{q}_t = \zeta \hat{y}_0 e^{-\lambda t} \tag{19}$$

[10] Such log-linear expansions are common when characterizing transitional dynamics in macroeconomic models. Specifically, the system's structure is similar to that of the neoclassical growth model in which the equilibrium is characterized by a dynamic system with one state and one control – capital and consumption – and two equations – a forward looking Euler equation and backward looking capital accumulation equation.

where

$$\lambda = -\frac{\rho}{2} + \sqrt{\left(\frac{\rho}{2}\right)^2 + \delta\left((\rho+\delta) + \frac{\psi}{\theta+\kappa\eta}\right)} \quad \textit{and} \quad \zeta = -\frac{1}{\theta-1}\frac{1}{\rho+\delta+\lambda}$$

Proof. See Appendix A.3. ■

Proposition 1 shows that the equilibrium always exists and is unique, which is a consequence of saddle-path stability. Conditional on the overall adjustment in relative output $\hat{y}_0$, the relative output y_t and the present discounted value of the relative wage q_t have a constant convergence rate of λ to the stationary equilibrium. The system immediately yields the path of the the relative wage along the transition since $\hat{\omega}_t = \frac{1}{1-\theta}\hat{y}_t$. Whenever $\theta > 1$, the changes in relative output and relative wage in the H-technology are negatively correlated along the transition. Intuitively, the endogenous evolution of the skill distribution shifts the relative labor supply in high-tech production, which generates an increase in the relative high-tech supply and, consequently, a decrease in the relative wage $\hat{\omega}_t$ and its presented discounted value $\hat{q}_t$.

The system describing the evolution of $\{\hat{q}_t, \hat{y}_t\}$ is isomorphic to the q-theory of capital investment (Tobin (1969), Hayashi (1982)) where, in our model, $\hat{q}_t$ is the present discounted value of relative price of high-tech output. In other words, it is the shadow price of the human capital "asset" associated with having one additional unit of high-tech output. Whenever this marginal value is higher, the incentives to invest in high-i skills are stronger, which has a positive impact on the relative value of high-tech output $\hat{y}_t$. As in the q-theory, the elasticity of relative output value $\hat{y}_t$ with respect to $\hat{q}_t$ depends on parameters governing the costs of adjustment (i.e, the term $\delta\psi$). However, as opposed to the traditional q-theory where all units of the capital stock are homogenous and perfectly substitutable, our model features imperfect substitution of human capital across technologies. Thus, the elasticity of y_t to q_t also depends on the degree of technology-skill specificity measured by η.

The connection between our model and the q-theory of investment becomes clear when we analyze the evolution of the skill distribution along the equilibrium path. In particular, Corollary 1 characterizes the transitional dynamics of the allocation of workers across technologies, which is summarized by the transitional dynamics of the skill distribution $s_t(i)$ and the assignment threshold l_t.

Corollary 1 (Worker allocations) *Given $\hat{s}_0(i)$ and the corresponding $\hat{y}_t, \hat{q}_t$, the evolution of the skill-distribution $\hat{s}_t(i)$, the optimal lotteries $\hat{s}_t(i)$, and the assignment threshold $\hat{l}_t$ are approximated* *by:*

$$\hat{s}_t(i) = \hat{s}_0(i)e^{-\delta t} + \int_0^t e^{\delta(\tau - t)}\hat{\hat{s}}_\tau(i)d\tau \tag{20}$$

$$\hat{\hat{s}}_t(i) = \left(\mathbb{I}_{i>l} - \int_l^1 s(i)di\right)\psi\hat{q}_t + o_t(i) \tag{21}$$

$$\hat{l}_t = \frac{\eta}{\theta - 1}\hat{y}_t \tag{22}$$

where $o_t(i)$ is such that $\int s(i)o_t(i)di = 0$.

Proof. See Appendix A.4 ■

The change in the optimally skill lottery along the transition depends centrally on the evolution of the relative return of skills employed in the high-tech production. In particular, changes in the present discounted value of relative prices $\hat{q}_t$ have a positive impact on types employed in the H-technology in the new stationary equilibrium ($i > l$). The parameter ψ controls the sensitivity of the optimal skill investment to such changes. The overall skill-distribution is then simply a population-weighted average of the skill-distributions of each generation. Since generations are born and die at rate δ, the population share at time t of the initial generation is $e^{-\delta t}$ whereas entering generation τ has a weight $\delta e^{\delta(\tau - t)}$. Finally, the assignment threshold changes are driven by changes in relative prices $\hat{\omega}_t = \frac{1}{\theta-1}\hat{y}_t$. The sensitivity to such changes is given by η.

3 The Economy's Adjustment to a Skill-biased Technological Innovation

We now analyze the dynamic adjustment of our economy to one-time permanent increase in the relative productivity A. Because this innovation increases the relative productivity of workers with higher skill-types i that are sorted into the H sector, we refer to it as a skill-biased technological innovation. First, we use the results above to characterize in closed-form the joint dynamics of q_t and y_t and worker allocations across technologies. Second, we analyze how the parameters governing the short-run and long-run skill supply affect different features of these dynamics. Third, we show the impact of the shock on average welfare and lifetime inequality for different worker generations, connecting them to the dynamics of q_t and y_t. Finally, we show that ignoring the adjustment across generations will typically lead to incorrect conclusions regarding average welfare and lifetime inequality because it confounds short-run and long-run skill supply elasticities.

3.1 Dynamic responses of labor market outcomes

We analyze the economy's adjustment following an unanticipated increase in the productivity of high-tech production at time $t = 0$, $\Delta log(A)$. We assume that immediately prior to the shock at time $t = 0^-$, the economy of Section 2 is in a stationary equilibrium with skill distribution given by $s_0(i)$. The shock only augments the productivity of skill types employed in high-tech production. Given the positive sorting of types into the H technology, the innovation is biased towards high-i types.

We start by characterizing the dynamic responses of $log(q_t)$ and $log(y_t)$ following the technological shock. We denote the change in outcomes from their initial levels as $\Delta log(q_t) \equiv log\left(q_t/q_{0^-}\right)$ and $\Delta log(y_t) \equiv log\left(y_t/y_{0^-}\right)$.

Proposition 2 (Dynamic responses) *Given a skill-biased technological innovation $\Delta log(A)$, the dynamic responses $\Delta log(y_t)$ and $\Delta log(q_t)$ are approximated by:*

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

where $\chi \equiv \left(\theta+\kappa\eta+\frac{\psi}{\rho+\delta}\right)(\rho+\delta)$.

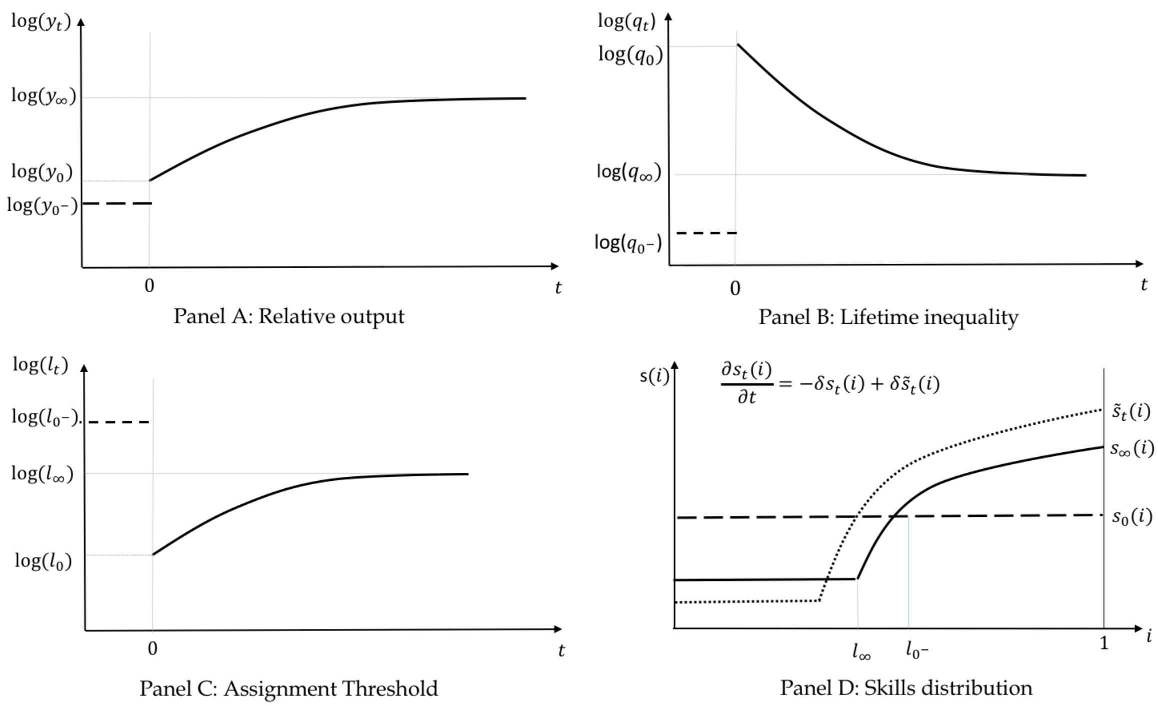
Proof. See Appendix A.5. ■

The proposition shows that when $\theta > 1$ both y_t and q_t increase on impact and in the long-run. However, along the transition, q_t falls while y_t increases at rate λ. To derive these expressions, first note that the results in Proposition 1 immediately yielded the transitional dynamics of y_t and q_t given an initial condition for $\hat{y}_0 = \log(y_0/y_\infty)$. Then, to characterize the full dynamic responses, in Appendix A.5 we simply derive expressions for the short-run and long-run changes in y_t implied by the technological shock, $\Delta log(A)$, given the economy's initial skill distribution $s_0(i)$. Furthermore, the expressions highlight how both the short-run and long-run skill supply parameters, η and ψ, crucially shape the dynamic responses—a point we return to in Section 3.2 when we show comparative statics with respect to these parameters.

We are now ready to provide a unified account of the economy's adjustment to a skill-biased technological shock by combining this proposition with the allocations of different generations of workers across technologies we characterized in Corollary 1. Figure 2 provides a graphical illustration.

The productivity increase leads to an increase in the demand for workers in the high-tech sector. In equilibrium, the relative price increases on impact ($\Delta log(\omega_0) > 0$). This causes workers of the old generation that had relatively low skill-types to enter the high-tech sector ($\Delta log(l_0) < 0$). As in static Roy models, the extent to which they do so crucially depends on how substitutable these workers are at the margin, i.e., the short-run skill-supply elasticity η. When $\theta > 1$, both the reallocation of workers and the direct effect of the productivity increase lead to higher relative output value on impact ($\Delta log(y_0) > 0$). Along the transition, younger entering workers decide to invest in high-i skills that are complementary to high-tech production in anticipation of higher relative wages (as captured by $\Delta log(q_t) > 0$). Thus,

Figure 2: The economy's adjustment to a skill-biased technological shock



they enter the economy with a skill-distribution that has a larger mass of high-i skill types compared to older generations. The extent to which they do so is determined by ψ because, as discussed in the previous section, it governs the sensitivity of the skill lottery to changes in q_t. Then, the overall skill-distribution changes as older generations are replaced with younger generations at rate δ. This increase in the mass of high-i workers expands the supply of high-tech output over time ($\Delta log(y_t) > 0$), triggering a decline in the current and present discounted value of relative wages ω_t, q_t and an increase in the assignment threshold l_t. This implies that intermediate-i types that initially entered the high-tech sector are displaced over time. Finally, compared to the initial equilibrium of the economy, the new long-run equilibrium entails a higher relative wage and output, and a larger mass of workers in the high-tech sector (driven both by a stationary skill-distribution with higher mass in high-i types and the lower assignment threshold).

3.2 Comparative Statics: short-run and long-run skill supply elasticities

We now turn to a detailed investigation of how the parameters that govern the short-run and long-run skill supply elasticities affect the economy's adjustment to the skill-biased technological innovation. For any given skill distribution, the parameter η controls the short-run skill supply elasticity to changes in the relative wage. In contrast, the parameter ψ controls the long-run skill supply elasticity to changes in the present discounted value of the relative wage in. These are the two main ingredients of the model governing the worker's ability to adjust to the shock and determine the transitional dynamics of labor market outcomes and welfare.

Short-run skill supply elasticity. Consider first how the technology-skill specificity affects the impulse response function of labor market outcomes. Since this parameter is intrinsically related to the extent of skill-technology specificity, this comparative static exercise can be thought as a comparison of the transitional dynamics following shocks to technologies with different degrees of skill specificity. For instance, an economy with a higher η has a lower degree of the technology-skill specificity in high-tech production and, therefore, a higher elasticity of relative employment in high-tech production.

The following proposition formally establishes how η affects the impulse response functions of the economy.

Proposition 3 (Comparative statics with respect to η)

1. *Short-run adjustment*

$$\frac{\partial \Delta log(y_0)}{\partial \eta} > 0 > \frac{\partial \Delta log(q_0)}{\partial \eta}$$

2. *Long-run adjustment*

$$\frac{\partial \Delta log(y_\infty)}{\partial \eta} > 0 > \frac{\partial \Delta log(q_\infty)}{\partial \eta}$$

3. *Persistence*

$$\frac{\partial \left(\int_0^\infty |\hat{y}_t| \, dt \right)}{\partial \eta} < 0, \quad \frac{\partial \left(\int_0^\infty \hat{q}_t dt \right)}{\partial \eta} < 0$$

Proof. See Appendix A.6. ■

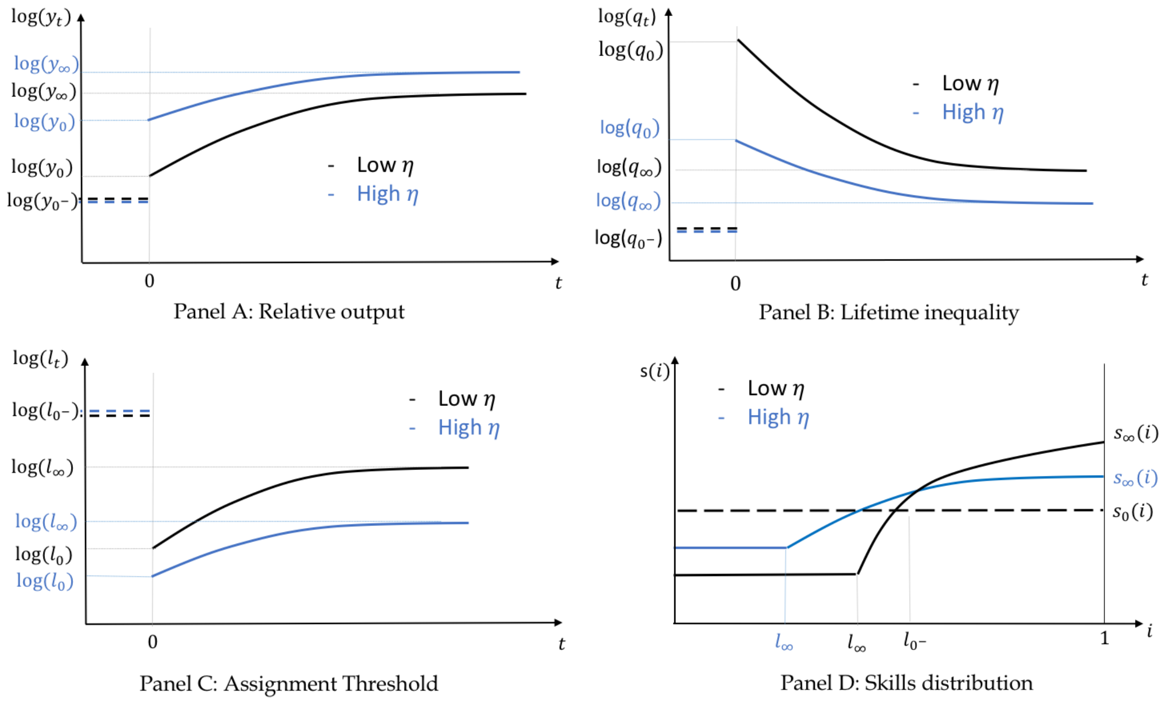
To fix ideas, Figure 3 illustrates the results in Proposition 3 with the impulse response functions of two economies: the blue lines show the responses of an economy with a high value of η, and the black lines show the responses of an economy with a low value of η.

In the short run, the economy with a higher η exhibits a larger change in the relative output of high-tech production and a lower change in the relative lifetime wage in high-tech production. This follows from the fact that, in the economy with a high η, it is easier to reallocate workers across technologies in response to the shock, implying that

$$\frac{\partial |\Delta log(l_0)|}{\partial \eta} > 0 \quad \text{and} \quad \frac{\partial \Delta \log(\omega_0)}{\partial \eta} < 0.$$

As a consequence of the smaller change in the relative lifetime wage in high-tech production when η is higher, younger workers have weaker incentives to invest in new skills. Thus, as we can see from Corollary 1, the overall change in the skill distribution implied

Figure 3: Comparative statics with respect to η



Panel A: Relative output

Panel B: Lifetime inequality

Panel C: Assignment Threshold

Panel D: Skills distribution

by the shock is smaller as well. This then implies that the transitional dynamics from the short- to the long-run are less persistent, as measured by the cumulative impulse response (e.g., $\int_0^\infty \hat{q}_t dt$). Finally, while the smaller change in the skill-distribution could have implied a smaller (larger) overall long-run adjustment in relative output (lifetime inequality), it turns that that larger (smaller) short-run response dominates. Thus the long-run adjustment in relative output (lifetime inequality) is larger (smaller). Taken together, these results show that the high η economy front-loads the adjustment to the technological innovation.

Long-run skill supply elasticity. We now consider how the parameter ψ affects the economy's adjustment to the technological innovation in high-tech production. This comparative statics exercise illustrates how economies with different costs of skill investment—and thus different degrees of long-run skill supply elasticity—respond to skill-biased technological shocks over time.

Proposition 4 (Comparative statics with respect to ψ)

1. *Short-run adjustment*

$$\frac{\partial \Delta log(y_0)}{\partial \psi} = 0 > \frac{\partial \Delta log(q_0)}{\partial \psi}$$

2. *Long-run adjustment*

$$\frac{\partial \Delta log(y_\infty)}{\partial \psi} > 0 > \frac{\partial \Delta log(q_\infty)}{\partial \psi}$$

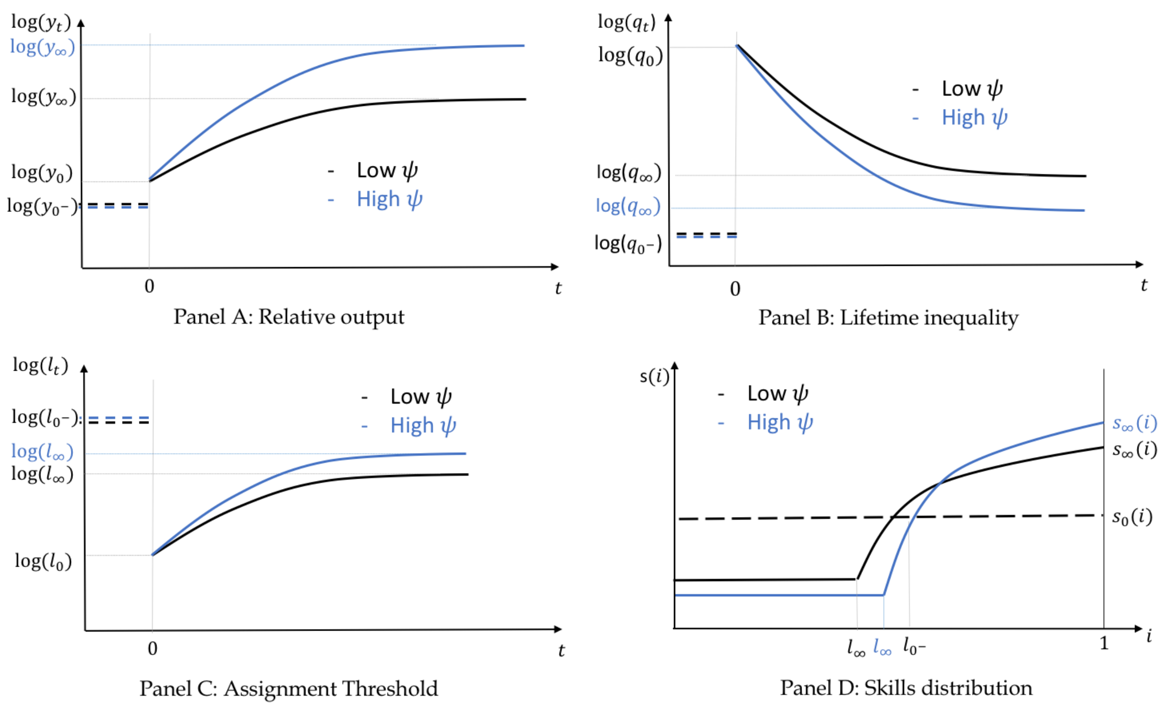
3. *Persistence*

$$\left.\frac{\partial \left(\int_0^\infty |\hat{y}_t|\, dt\right)}{\partial \psi}\right|_{\psi=0} > 0, \quad \left.\frac{\partial \left(\int_0^\infty \hat{q}_t dt\right)}{\partial \psi}\right|_{\psi=0} > 0$$

Proof. See Appendix A.6. ■

Again, we illustrate the results in Proposition 4 with a graphical representation of the impulse response functions of two economies in Figure 4. The blue lines depict the adjustment of an economy with a high value of ψ, and the black lines represent the responses of an economy with a low value of ψ. Accordingly, the "black" economy is closer to a static model with inelastic skill supply, while the "blue economy" entails dynamics arising from the ability of younger generations of workers to invest in skills.

Figure 4: Comparative statics with respect to ψ



The first part of Proposition 4 indicates that the short-run response of relative output is identical in both economies. This is because the parameter ψ does not affect the relative employment supply of old generation borns before the shock, neither the change in relative wages $\Delta log(\omega_0)$. However, a higher ψ decreases the short-run lifetime inequality increase after the shock because future relative wages fall by more as consequence of the larger increase in the future supply of skills (which is evidenced in Corollary 1). The later also implies that relative output (lifetime inequality) increases more (less) in the long-run. Finally, because there is larger change in the skill-distribution along the transition, the persistence of both lifetime inequality and relative output increases when ψ is higher. Taken together, these results show that the high ψ economy back-loads the adjustment to the technological innovation.

3.3 Welfare analysis

The average welfare of young workers born after the shock of generation τ is:

$$U_\tau = \int_0^1 \tilde{s}_\tau(i) V_\tau(i) di - \frac{1}{\psi} \int_0^1 \tilde{s}_\tau(i) log\left(\frac{\tilde{s}_\tau(i)}{\bar{s}(i)}\right) di$$

where, as a reminder, $V_\tau(i)$ is the present discounted value of log-wages

$$V_\tau(i) = \frac{log(\alpha(i))}{\rho+\delta} + log\left(Q_\tau(i)\right) - \int_\tau^\infty e^{-(\rho+\delta)t} log\left(P_t\right) dt$$

which depends on $log(Q_\tau(i))$ which captures the present-discounted value of log-relative wages $log(\omega_t)$ and the present discounted value of the ideal price index $P_t = (1+y_t)^{1-\theta}$.

Furthermore, because of the envelope theorem, we have that that the average welfare of old workers born before the shock who cannot invest skills is (to a first order approximation) equal to U_0, i.e., the average welfare of the initial young generation born after the shock who can invest in skills.

Then, consider generation welfare-weights $re^{-r\tau}$ that discount the welfare of generations far in the future. We can then compute the average welfare across all generations as:

$$\bar{U} = r \int_0^\infty e^{-r\tau} U_\tau d\tau$$

Analogously, we can compute the average lifetime welfare inequality as:

$$\bar{\Omega} = r \int_0^\infty e^{-r\tau} log(q_\tau) d\tau$$

The following proposition shows how the change in average welfare $\Delta\bar{U} \equiv \bar{U} - U_{0^-}$ and lifetime inequality $\Delta\bar{\Omega} \equiv \bar{\Omega} - log(q_{0^-})$ depend on the long-run response and persistence of $log(q_t)$ that we characterized in Propositions 3 and 4.[11]

Proposition 5 (Average welfare and lifetime welfare inequality) *The change in average wel-*

[11]Note that, because of log-utility in consumption, $(\rho+\delta)\Delta\bar{U}$ has the interpretation of a consumption equivalent variation.

fare $\Delta\bar{U}$ and lifetime inequality $\Delta\bar{\Omega}$ are approximately:

$$\Delta\bar{U} = \frac{1}{\rho+\delta}\frac{y_\infty}{1+y_\infty}\Delta log(A) - \left(\frac{y_\infty}{1+y_\infty} - \int_l^1 s(i)di\right)\Delta\bar{\Omega}$$
$$\Delta\bar{\Omega} = \Delta log(q_\infty) + \frac{\lambda r}{r+\lambda}\int_0^\infty \hat{q}_\tau d\tau$$

Proof. See Appendix A.7 ■

Regarding $\Delta\bar{U}$, the term $\frac{1}{\rho+\delta}\frac{y_\infty}{1+y_\infty}\Delta log(A)$ is simply the direct effect of the skill-biased innovation on the ideal price index, which is lower at all times after the shock and thus increases average real wages and welfare. The second term is a combination of two effects. First is the positive effect on welfare due to the increase in the present discounted value of relative wages for those workers that are in the high-tech sector ($\int_l^1 s(i)di\Delta\bar{\Omega}$). Second is the negative effect on average welfare of all workers due to a smaller decline in the ideal price index when relative wages increase by more, which is the flipside of a smaller expansion in the supply of goods ($\frac{y_\infty}{1+y_\infty}\Delta\bar{\Omega}$). Then, for a given direct effect in the first term, an increase in average lifetime inequality decreases average welfare if and only if the share of output in the high-tech sector ($\frac{y_\infty}{1+y_\infty}$) is larger than the share of employment in the high-tech sector ($\int_l^1 s(i)di$) and thus the effect of the ideal price index for all workers dominates the effect of relative wages for workers in the high-tech sector.

Regarding $\Delta\bar{\Omega}$, it is higher when either the long-run change or the persistence in lifetime inequality are higher. Furthermore, when $0 < r < \infty$, it is also higher when λ is higher (conditional on the long-run response and the persistence). This is because λ govern how fast lifetime inequality decays along the transition, and as such, the differences *between* generations in lifetime inequality. When λ is higher, generations far in the future have similar lifetime inequality than, for example, the initial generation. Then, because such future generations are discounted at rate r, the average lifetime inequality across generations increases when $\frac{\lambda}{r+\lambda}$ is higher (conditional on the long-run response and the persistence). Relatedly, when the discount rate r is lower, average lifetime welfare inequality is lower. To see why, consider the case with $r \to 0$ where we are only giving a positive weight to generations born at $\tau \to \infty$. Since lifetime inequality decays along the transition, such generations have lower lifetime inequality compared to generations present around the time the innovation occurred. The opposite is true when we consider the case $r \to \infty$ and we only give a positive weight to the initial generations.

3.4 What do we miss by ignoring the adjustment across generations?

The short-run changes approach. A common approach in empirical macroeconomics, labor and trade is to identify the responses of labor market variables to a technological or tariff shock over some fixed time horizon. The time horizon is often short (typically ranging from one to 20 years) due to data limitations and the difficulty of finding plausibly exogenous variation over longer horizons. We next show how this approach also leads to incorrect predictions regarding the consequences of skill-biased technological innovations. In particular, it understates the average welfare increase and overstates the lifetime welfare inequality increase resulting from the innovation. This is because it only considers short-run responses which depend on the short-run skill supply elasticity but ignores the future adjustment across generations. Yet, we show in the next section how this bias can be solved by looking at the responses of young and old separately over a short time horizon following a technological innovation.

We begin by noting that we can write:

$$\Delta\bar{\Omega} = \frac{\Delta log(\omega_0)}{\rho+\delta} - \left(\frac{\lambda}{\rho+\delta} + \frac{\lambda}{r+\lambda}\right)\lambda\int_0^\infty \hat{q}_\tau$$
$$\Delta log(\omega_0) = \frac{(\theta-1)\Delta log(A)}{\theta+\kappa\eta}$$

By looking at the denominator in $\Delta log(\omega_0)$, we see that the present discounted value of the short-run change in wage inequality $\frac{\Delta log(\omega_0)}{\rho+\delta}$ depends only on η and not ψ.

Next, consider parameterizations with the same η but different ψ's. In particular, the case of $\psi = 0$ is a special case where there is no adjustment across generations and our model reduces to a static Roy model. From Proposition 5, it is straightforward to derive that,

$$\Delta\bar{\Omega}|_{\eta,\psi>0} - \Delta\bar{\Omega}|_{\eta,\psi=0} = -\left(\frac{\lambda}{\rho+\delta} + \frac{\lambda}{r+\lambda}\right)\lambda\int_0^\infty \hat{q}_\tau < 0$$

The above implies that a researcher using a static Roy model to compute welfare using observed short-run responses will overstate the lifetime welfare inequality increase and understate the average welfare increase (if $(\frac{y_\infty}{1+y_\infty} > \int_l^1 s(i)di)$ following a skill-biased technological innovation. Because relative wages fall along the transition once the adjustment across generations is taken into consideration, the short-run change in relative wages overstates both the short-run change in lifetime inequality for the initial generation as well as the future decline for future generations. Thus, given the same short-run change which only depends on η, the increase in average lifetime welfare inequality across generations $\Delta\bar{\Omega}$ is smaller when $\psi > 0$. As a result, a bias arises using observed short-run responses and a static Roy model to analyze how economies adjust to technological innovations.

The long-run changes approach. A common approach in quantitative macroeconomics and trade is to use static assignment models to study the consequences of long-run changes in technology and tariffs.[12] We next show how this approach leads to incorrect predictions regarding the consequences of skill-biased technological innovations. In particular, it overstates the average welfare increase and understates the lifetime welfare inequality increase

[12] Or sometimes use dynamic models but only look at changes across steady states.

resulting from the innovation. This is because, by ignoring the adjustment across generations, it confuses short and long-run skill supply elasticities which have similar implications for long-run responses but different implications for the adjustment along the transition (as we have showed before).

We begin by reminding that:

$$\Delta log(q_\infty) = \frac{\frac{1}{\rho+\delta}(\theta-1)\Delta log(A)}{\theta+\eta\kappa+\frac{\psi}{\rho+\delta}}$$

By looking at the denominator, we see that any given long-run change in lifetime inequality can be obtained with either a low η and high ψ or viceversa. In particular, the case of $\psi = 0$ is a special case where there is no adjustment across generations and our model reduces to a static Roy model.

Next, consider different combinations of η and ψ such that they lie in the lochus $\chi(\psi,\eta) \equiv \theta+\eta\kappa+\frac{\psi}{\rho+\delta} = \bar{\chi}$ and thus achieve the exact same long-run change in lifetime inequality. From Proposition 5, it is straightfoward to see that, given parameterization $\psi^a = 0, \eta^a > 0$ and $\psi^b > \psi^a, \eta^b < \eta^a$ such that both parameterizations lie in the above lochus,

$$\Delta\bar{\Omega}^b - \Delta\bar{\Omega}^a = \frac{\lambda^b r}{r+\lambda^b}\int_0^\infty \hat{q}^b_\tau d\tau > 0$$

The above implies that a researcher using a static Roy model will understate the lifetime welfare inequality increase and overstate the average welfare increase (if $(\frac{y_\infty}{1+y_\infty} > \int_l^1 s(i)di)$ following a skill-biased technological innovation. The reason is that she will interpret the combined reduced-form long-run elasticity in χ as a long-run skill supply elasticity in her static model. However, because different combinations of short and long-run elasticities that achieve the same combined reduced-form long-run elasticity lead to different implications for both the persistence in lifetime welfare inequality and λ, they then lead to different implications for $\Delta\bar{\Omega}$ and $\Delta\bar{U}$. In particular, when we consider our model with $\psi = 0$ and a high η to match a given long-run reduced-form elasticity, then the adjustment in lifetime welfare inequality to its new long-run level is instantaneous. However, when $\psi > 0$ and η is lower, then the adjustment in the skill-distributon across generations is slow. As we saw before, in this case lifetime welfare inequality overshoots its long-run level and then slowly falls along the transition. Thus, given the same long-run change, the increase in average lifetime welfare inequality across generations $\Delta\bar{\Omega}$ is higher when $\psi > 0$. As a result, a bias arises when ignoring the adjustment across generations and using a static Roy model to analyze how economies adjust to technological innovations.

4 Observable Responses in Labor Market Outcomes Between and Within Generations

Our theoretical results establish the distinct roles of the short- and long-run skill supply elasticity in shaping the economy's adjustment path following skill-biased technological shocks. In this section, we show that these two channels also have different observable predictions for the evolution of labor market outcomes within and between worker generations. In particular, in response to a technolgical shock, we connect the parameters of short- and long-run skill supply elasticity respectively to the employment adjustment of old and young generations.

From Proposition 2, we obtain the first observed prediction of the model: the evolution of relative payroll in high-tech production,

$$\Delta \log y_t = \left(\frac{1+\kappa\eta}{\theta+\kappa\eta} + \frac{\psi}{\chi} \frac{\theta-1}{\theta+\kappa\eta} (1-e^{-\lambda t}) \right) (\theta-1) \Delta \log(A). \tag{23}$$

The dynamic response in (23) summarizes two important observable implications of the model. First, it shows that, whenever $\theta > 1$, a positive shock in high-tech productivity yields a proportional increase in the relative value of high-tech production. Second, it shows that relative high-tech output grows along the transition at rate λ.

In Section 3, we also characterize the evolution of the present value of relative high-tech wage q_t, as well as the adjustment of the skill distribution, $s_t(i)$, and the skill-technology assignment, l_t. There are two challenges to empirically evaluate the model's predictions regarding these outcomes.

First, data on these outcomes is not readily available for researchers. As in the q-theory of capital investment, q_t is a forward-looking variable whose measurement requires knowledge of the entire path of relative wages. In addition, without taking a explicit stance on what observable attributes determine worker skills in different activities, it is not possible to directly measure the skill distribution and the skill-technology assignment. Thus, we focus on observable labor market outcome: relative employment across occupations. To this end, we map high-tech production in our model to a set of occupations disproportionately augmented by a technological innovation. Through the lens of the model, the evolution of relative employment combine changes in both the skill-technology assignment and the skill distribution.

Second, it is often hard to exactly match technological innovations to specific dates in which workers start adjusting their skills. This task is even harder if we acknowledge that, in practice, young workers invest on skills prior to entering the labor force in the form of schooling, as well as after joining the labor form in the form of vocational training or on-the-job learning. We circumvent this problem by deriving changes in relative employment for young and old generations of workers. We define young generations as all workers born after a specific year prior to the shock. This implies that the group of young generations contains workers whose skill investment was made before and after the technological innovation.

Specifically, we consider the relative employment in high-tech production of two groups of workers: old workers born at period $t = -x$ and young workers born at period $t = -x$. In period $t \geq 0$, the relative high-tech employment of these worker groups are given by

$$e_t^{old} = \frac{\int_{l_t}^{1} s_0(i)di}{\int_0^{l_t} s_0(i)di} \quad \text{and} \quad e_t^{young} = \frac{\tilde{x}_0 e^{-\delta t} \int_{l_t}^{1} s_0(i)di + \delta \int_0^t e^{\delta(\tau - t)} \int_{l_t}^{1} \tilde{s}_\tau(i)did\tau}{\tilde{x}_0 e^{-\delta t} \int_0^{l_t} s_0(i)di + \delta \int_0^t e^{\delta(\tau - t)} \int_0^{l_t} \tilde{s}_\tau(i)did\tau}.$$

where $\tilde{x}_0 \equiv 1 - e^{-\delta x}$ is the share of the young generation in the economy's population before the shock (i.e., time $t = 0^-$).

For both worker groups, the skill-technology assignment is identical and determined by the threshold l_t. Notice that all workers of the old generations have the pre-shock skill distribution, $s_0(i)$. However, the skill distribution of young generations combines the pre-shock distribution $s_0(i)$ and the post-shock lotteries $\tilde{s}_\tau(i)$. The overlapping generation structure of the model implies that the relative share of workers in the young generation with the pre-shock skill distribution decays at the constant rate δ.[13]

We now show that, in the model, employment responses of old generations are mainly driven by the magnitude of skill-technology specificity (i.e., short-run skill supply elasticity), while employment responses of young generations are also driven by the magnitude of the skill investment cost (i.e., long-run skill supply elasticity).

Relative employment of old generation: Short-run skill supply elasticity. In Appendix A.8, we show that the change in the relative employment of old generations can be approximated as

$$\Delta \log e_t^{old} \approx \frac{\eta}{\theta + \kappa\eta} \frac{1}{e_{H,0^-}} \left(1 - \frac{\psi}{\chi}(1 - e^{-\lambda t})\right) (\theta - 1)\Delta \log A. \tag{24}$$

where $e_{H,0^-}$ is the high-tech employment share at $t = 0^-$.

Among old generations, the increase in the relative productivity of high-tech production induces the reallocation of older workers towards high-tech production whenever $\theta > 1$. The expression indicates that this positive effect on the relative high-tech employment becomes weaker over time. As discussed in the previous section, this follows from the expansion of high i skills in the young generations, which displaces from high-tech production old workers with marginal skills – i.e., those with skills $i \in (l_0, l)$.

Expression (24) shows that the magnitude of the increase in high-tech relative employment is increasing in the short-run skill supply elasticity, η. To see this more clearly, consider

[13] We assume that workers born before the shock in the young generation do not adjust their skills. It is possible to allow part of these workers to adjust their skills when the shock occurs at $t = 0$. In this case, rather than $s_0(i)$, the skill distribution of these workers would be a mix of $s_0(i)$ and $\tilde{s}_0(i)$. This extension does not alter the main insights of the model, but it reduces the short-to-long adjustment in relative output $\hat{y}_0$.

the relative employment response of old workers at impact:

$$\frac{\Delta \log e_0^{old}}{\Delta \log A} \approx \frac{\eta}{\theta + \kappa\eta} \frac{\theta - 1}{e_{H,0^-}}.$$

This expression indicates that, conditional on the shock size, a higher η induces larger changes in relative employment in the short-run. Whenever η is small, old workers do not change their relative employment following the shock. Intuitively, this parameter controls the skill-technology specificity and, therefore, how much changes in the relative high-tech wage affect the assignment threshold and relative employment.

Relative employment of young generation: Long-run skill supply elasticity. Turning to the employment response among young generations, Appendix A.8 also establishes that

$$\Delta \log e_t^{young} \approx \Delta \log e_t^{old} + \frac{\psi}{\chi} \frac{1 - e^{-\lambda t}}{1 - (1 - \tilde{x}_0) e^{-\delta t}} (\theta - 1) \Delta \log A. \tag{25}$$

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

Expression (24) shows that the between-generation difference in relative high-tech employment is increasing in the parameter of long-run skill supply elasticity, ψ. To see this more clearly, consider the between-generation employment difference in the long-run:

$$\frac{\Delta \log e_\infty^{young} - \Delta \log e_\infty^{old}}{\Delta \log A} \approx \frac{\psi}{\psi + (\theta + \kappa\eta)(\rho + \delta)} (\theta - 1).$$

Conditional on the shock size, a higher ψ yields stronger employment differences across generations in the long-run. Intuitively, this parameter controls the sensitivity of the skill supply of incoming cohorts to relative lifetime earnings. A higher elasticity implies that young workers adjust more intensively their skills, amplifying differences between their relative high-tech employment that of older generations.

The next section relies on the insights obtained from (23)–(25) to empirically document the importance of short- and long-run skill supply elasticity for the dynamic adjustment of labor markets following the arrival of new technologies. In Section 6, we combine the estimated empirical responses and expressions (23)–(25) to calibrate the model's parameters and quantify the welfare gains of the new technologies.

5 Technological Transitions in Germany: Adjustment Across Generations

We now use the results of the previous section to empirically evaluate the importance of accounting for both short-run and long-run skill supply elasticities when analyzing the economy's dynamic adjustment to skill-biased technological shocks. Specifically, this section studies how the German labor market adjusted to recent technological innovations like computers and internet. Since usage of these new technologies is strongly correlated to an occupation's intensity in cognitive tasks, we take cognitive-intensive occupations to be the set of economic activities disproportionately augmented by the shock. At the national level, between 1997 and 2014, we document that cognitive-intensive employment increased among young generations, but did not change among old generations. We then exploit quasi-exogenous cross-regional variation in the adoption timing of broadband internet to estimate the impulse response functions of employment and payroll across occupations and cohorts. We find that, in early adopting regions, relative employment and output in cognitive-intensive occupations increase after 2005, but do not change in the pre-shock period of 1997-2004. These responses are almost entirely driven by the reallocation of young generations of workers. Our results suggest that skill supply elasticity is very low in the short-run, but it is positive in the long-run.

5.1 Data

Our main source of information on German labor market outcomes is the LIAB Longitudinal Model. This is a linked employer-employee dataset released by the Institute for Employment Research (IAB). The dataset includes individual information on total earnings and days worked, as well as data on education, occupation, full-time status, and employer location. Our analysis focuses on full-time males aged 20–60 residing in West Germany. We construct our main sample of employed individuals following closely the procedure in Card, Heining, and Kline (2013).

In our empirical application, we evaluate the impact of technological shocks on employment and payroll across cohorts, occupations, and regions.[14] We define as regional markets the 323 administrative districts in West Germany. This measure of local labor market in Germany has been used by Dauth, Findeisen, and Suedekum (2014) and Huber (2018). Our dataset contains information on the district where each establishment is located in 1999-2014. Whenever available, we use the worker's establishment's district in 1999 to construct worker district affiliation in 1997-1998.

We link each employed individual to one of the 120 occupations in the LIAB dataset. We use the BERUFNET dataset to define each occupation's cognitive intensity as the share of

[14] Figure 9 in Appendix B shows that, in our baseline sample, the between-component of log-wage variance associated with cohorts-occupation-region triples explains around 60% of the substantial rise in German inequality in the 2000s. This component has a similar magnitude as the between-firm log-wage variance emphasized by Card, Heining, and Kline (2013)

time spent on tasks that intensively require analytical non-routine and interactive skills.[15] We also obtain information about the daily activities performed by individuals employed in a subset of 85 occupations in the Qualification and Working Conditions Survey.

5.2 Cognitive-Intensity and Use of New Technologies Across Occupations

As a starting point, we analyze the types of tasks required by cognitive-intensive occupations. Figure 5 reports the correlation between the occupation's intensity in cognitive skills and the share of individuals in that occupation reporting to intensely perform each of the listed task. The top tasks performed in cognitive-intensive occupations are directly related to technological innovations recently introduced in the workplace: working with internet, in particular, and with computers, more generally. On the other extreme, individuals employed in the least cognitive-intensive occupations tend to perform routine tasks associated with manufacturing and repairing.

The results in Figure 5 are consistent with the evidence establishing the heterogeneous impact of new technologies on different tasks performed by workers – e.g., Autor, Levy, and Murnane (2003), Spitz-Oener (2006), Autor and Dorn (2013), and Akerman, Gaarder, and Mogstad (2015). Recent innovations, such as the computer and the internet, are more likely to be used in cognitive-intensive jobs whose daily activities require problem-solving, creativity, or complex interpersonal interactions. In contrast, these technological innovations lead to the substitution of routine-intensive jobs whose tasks follow well-understood procedures that can be codified in computer software, performed by machines or, alternatively, offshored over computer networks to foreign work sites. Thus, in the last decades, these types of skill-biased technogical innovations arguably raised the relative demand for cognitive tasks.

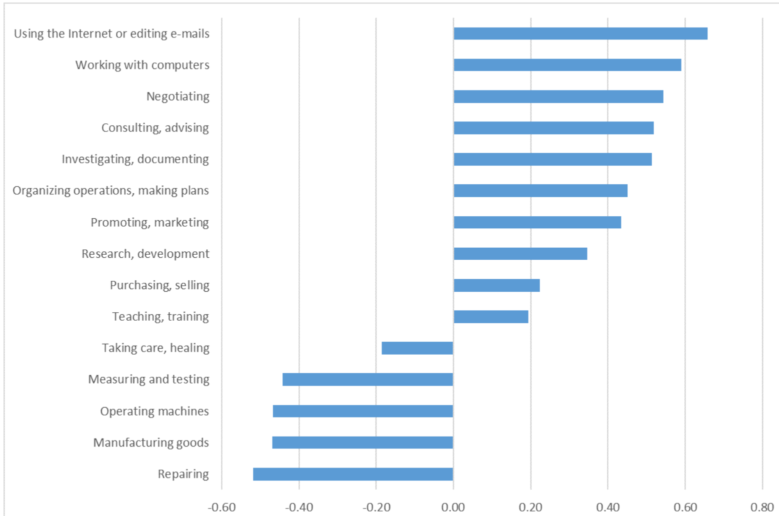
We then investigate whether these new technologies affected worker generations differently conditional on their occupation. We consider two generations: a young generation aged less than 40 years and an old generation aged more than 40 years.[16] Figure 6 shows that, while internet and computer usage are biased towards cognitive-intensive occupations, there were only small differences in the usage of these new technologies across worker cohorts employed in the same occupation in 2012. These results are complement the finding in Spitz-Oener (2006) that there were small between-cohort differences in the change of the task content of German occupations in the 1990s.

In the rest of this section, we take cognitive-intensive occupations to be the set of production activities being disproportionately augmented by recent technological innovations like the adoption of computers and internet in the workplace. Thus, in the analysis of Section 3, cognitive-intensive occupations correspond to the H technology whose productivity increases equally for all workers, independent of their age or skill. By pursuing this ap-

[15] The German Federal Employment Agency produces the BERUFNET dataset using expert knowledge about the skills required to perform the daily tasks in each occupation. We use the simple average in the occupation's cognitive intensity in the years of 2011-2013. We then collapse occupations into 100 quantiles based on their cognitive-intensity.

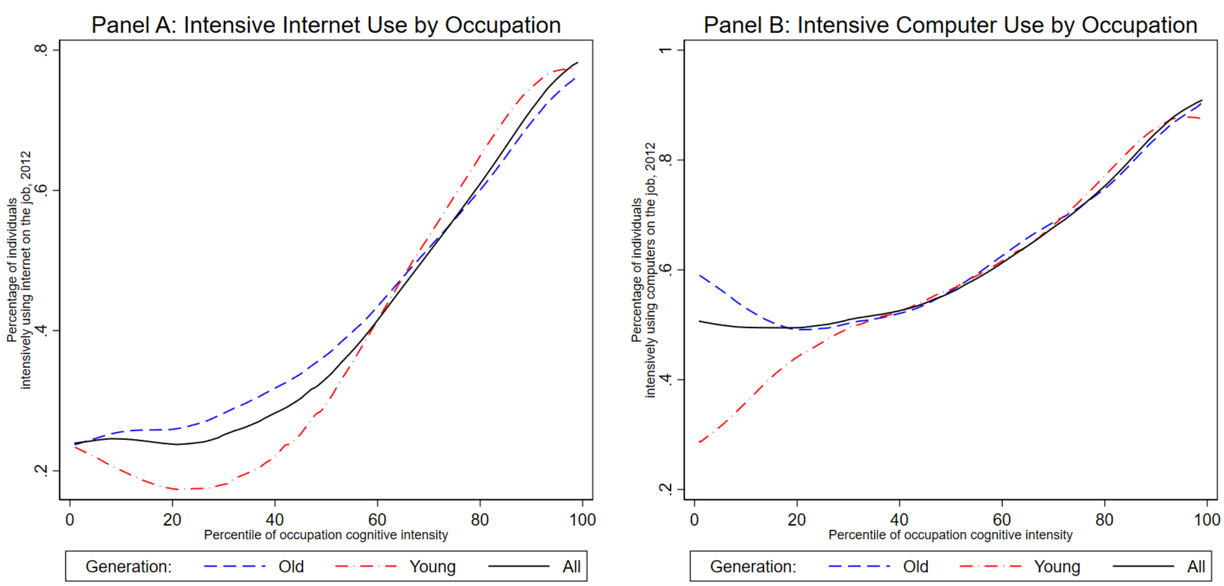
[16] Results are similar if we define young generations to include workers who are less than 30, 35 or 45 years old.

Figure 5: Cross-occupation correlation between cognitive intensity and perfomance of different tasks



Notes. Note. Sample of 85 occupations. The occupation task intensity is the share of individuals in that occupation reporting to intensively perform the task in the 2012 Qualification and Working Conditions Survey. The occupation cognitive-skill intensity is the share of time spent on cognitive-intensive tasks in the BERUFNET dataset (2011-2013).

Figure 6: Internet and Computer Usage by Occupation: Within- and Between-Generation



Note. Sample of 85 occupations in Working Condition Survey. For each occupation, we compute the share of individuals reporting to intensively use internet and computer on their job. Young generations defined as all workers born after 1960 and Old cohort as all workers born before 1960. The occupation cognitive-skill intensity is the share of time spent on cognitive-intensive tasks in the BERUFNET dataset (2011-2013). Figure reports the lowess smooth fit.

proach, we follow an extensive literature documenting how technological innovations affect jobs with a different task content —for a review, see Acemoglu and Autor (2011).

5.3 Cognitive-intensive Employment Growth in Germany

We now study how new technologies affected the employment composition in Germany. Motivated by the model's predictions in (24)–(25), we estimate generation-specific responses in the relative cognitive-intensive employment in Germany. Formally, for each worker generation g, we estimate the following linear regression:

$$\log Y^{g}_{o,2014} - \log Y^{g}_{o,1997} = \beta^{g}\bar{C}_{o} + X_{o}\gamma^{g} + \epsilon^{g}_{o} \tag{26}$$

where $Y^{g}_{o,t}$ is a labor market outcome in occupation o at year t of workers of generation g, $\bar{C}_o$ is the cognitive intensity of occupation o, and X_o is a control vector.[17] In this specification, the coefficient β^g captures the model's predictions outlined in Section 4: the differential employment growth in occupations with a higher cognitive intensity for the generation g.

Column (1) of Table 1 reports that the estimation of (26) for all generations of workers employed in 1997 and 2014. Panel A indicates that occupations with a higher cognitive intensity experienced stronger employment growth in the period. Compared to the least cognitive intensive occupation, the employment growth in the most cognitive intensive occupation was 0.9 log points higher. Panel B investigates the extent of polarization in employment growth across percentiles of the cognitive intensity of occupations. Employment growth was relatively stronger (weaker) in the occupations in top (bottom) of the cognitive intensity distribution. We obtain similar positive changes for the changes in the relative payroll of cognitive intensive occupations (see Table A2). Through the lens of the model, these results are consistent with the arrival of cognitive-biased technologies that shift the economy's employment composition towards cognitive-intensive occupations.

Columns (2) and (3) report the estimates of (26) for two worker generations: the old generation born before 1960 (column (2)) and the young generation born after 1960 (column (3)).[18] In the first year of our sample, the old generation is at least 37 years old and accounts for 40% of employment in Germany. Estimates indicate that old generations did not disproportionately enter cognitive-intensive occupations. This estimate is equivalent to the within-generation component of employment growth in cognitive-intensive occupations. In contrast, among young generations, we find statically significant stronger growth in occupations with a higher cognitive intensity. Accordingly, column (4) shows that the cognitive-intensive employment growth in Germany is almost entirely driven by its between-generation component – defined as the difference between the response for all generations in column (1) and for old generations in column (2).

[17] In all regressions, the control vector includes the 1997-2014 growth in the fraction of immigrants in the occupation, and occupational export exposure and import exposure (defined as the 1997-2014 growth in exports and imports in industries using the occupation, weighted by the share of that industry's employment in total occupational employment in 1997). These controls capture potential confounding effects on employment growth arising from the occupation's exposure to immigration and trade shocks affecting Germany during the period of analysis. Appendix Table A3 shows that results are similar without any controls and in the sample of native-born workers.

[18] Appendix Table A3 shows that results are qualitatively similar if we define young generations as workers born after 1955 or 1965. Results are also similar if we define young generations as those less than 40 years old in each year.

Table 1: Cognitive Intensity and Employment Growth in Germany

Dependent variable: Log-change in occupation employment, 1997-2014

	All (1)	Old (2)	Young (3)
Panel A: Linear specification			
Cognitive intensity	0.007**	0.000	0.011***
	(0.003)	(0.003)	(0.004)
Panel B: Nonlinear specification			
Percentile of cognitive intensity			
Low: below percentile 30	-0.262	-0.106	-0.345
	(0.271)	(0.283)	(0.290)
Medium: percentiles 30-60	0.086	0.214	0.113
	(0.209)	(0.222)	(0.231)
High: above percentile 60	0.463*	0.016	0.691**
	(0.242)	(0.225)	(0.231)

Notes. Sample of 120 occupations in 1997 and 2014. Young cohort defined as all workers born after 1960 and Old cohort as all workers born before 1960. All regressions include a set of baseline controls: growth in occupational exposure to exports from 1997-2014, growth in occupational exposure to imports from 1997-2014, and growth in the fraction of immigrants in the occupation in 1997-2014. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We can now connect the evidence in Table 1 to the two main features of our model. Result suggest that, following the arrival of new cognitive-biased technologies in Germany, there was stronger employment growth in cognitive-intensive occupations. The lack of employment reallocation for old generations suggests a low parameter of short-run skill supply elasticity. In our model, such a small response arises whenever skills are not easily transferable across occupations with different cognitive content. At the same time, the positive between-generation component of relative cognitive-intensive employment growth indicates a positive long-run skill supply elasticity that arises from the ability of young workers to adjust their skills in response to the technological innovation.

5.4 Dynamic Adjustment to the Broadband Internet Adoption

Having established these employment responses across occupations and generations in Germany, we now analyze the dynamic response to one cognitive-biased technological innovation: the introduction of broadband internet in the early 2000s. There are three reasons to focus on this particular innovation in Germany. First, the adoption of broadband internet was fast: the share of households using it increased from 0% in 2000 to over 90% in 2009. Second, it was spatially heterogeneous: across German districts in 2005, the mean share of household with broadband internet access was 76% and the standard deviation was 16%. Third, following Falck, Gold, and Heblich (2014), it is possible to isolate exogenous spatial variation in adoption timing coming from the suitability of pre-existing local telephone networks for broadband internet transmission.

These reasons imply that the introduction of broadband internet resembles a one-time permanent shock like the one studied in Section 3. Specifically, we rely on quasi-random cross-regional variation in the adoption of this technological innovation to estimate the impulse response functions of labor market outcomes across occupations and cohorts.[19]

5.4.1 Estimation Strategy

Our goal is to estimate the dynamic impact of broadband internet adoption in 2000-2007 on labor market outcomes across districts in Germany. For each year between 1997 and 2014, we estimate the following linear specification

$$Y_{io,t} - Y_{io,1999} = (\alpha_t + \beta_t \bar{C}_o)\, DSL_i + \delta_{o,t} + X_{io,t}\gamma_t + \epsilon_{io,t}, \tag{27}$$

where o denotes an occupation and i district in Germany. In this specification, $Y_{io,t}$ is a labor market outcome for occupation o of district i at year t (either employment or payroll), and DSL_i is the broadband internet penetration in district i in 2005 (normalized to have standard deviation of one). As before, $\bar{C}_o$ is the time-invariant measure of the cognitive-intensity of occupation o. The term $\delta_{o,t}$ is an occupation-year fixed-effect that absorbs any confounding shock that has the same impact on occupations and cohorts in all regions. Similarly, $X_{io,t}$ is a control vector which includes the dependent variable pretrend growth in 1995-1999 and initial district demographic characteristics. These controls account for differential growth in cognitive-intensive occupations in regions with different demographic characteristics.[20]

To obtain within- and between-generation responses, we also estimate a similar specifications for different worker generations g:

$$Y^g_{io,t} - Y^g_{io,1999} = \sum_{c \in \{\text{young, old}\}} (\alpha^c_t + \beta^c_t \bar{C}_o)\, 1_{[g=c]} DSL_i + \delta_{o,t} + \zeta_{g,t} + X_{io,t}\gamma_t + \epsilon^g_{io,t}, \tag{28}$$

where $Y^g_{io,t}$ is a labor market outcome for individuals of cohort g employed in occupation o of district i at year t. As above, we consider two generations: the old generation born before 1960 and the young generation born after 1960.[21] Notice that this specification also includes generation-year fixed-effects that capture nationwide trends in employment of different worker cohorts.

[19] Nakamura and Steinsson (2017) argue that well-identified causal effects are "powerful diagnostic tools for distinguishing between important classes of models." This idea guides our empirical analysis. It provides evidence about the causal effect of one type of skill-biased technological shock, which we then confront against our model's central predictions.

[20] The demographic controls are the college graduate population share, the manufacturing employment share, the immigrant employment share, and the age composition of the labor force. Appendix Table A4 shows results are robust to varying or removing the control vector. We obtain similar results when we drop the baseline controls and the control for pre-trends. Freyaldenhoven, Hansen, and Shapiro (2018) discuss the importance of controlling for pretrends. See also Dix-Carneiro and Kovak (2017) for an application controlling for regional pre-trends.

[21] Appendix Table A4 shows that results are qualitatively similar if we define young generations as workers born after 1955 or 1965. Results are also similar if we define young generations as those less than 35, 40 or 45 years old in each year.

We are mainly interested on the impact of broadband internet expansion on the relative outcome of cognitive-intensive occupations: β_t in (27) for the all workers, and β_t^g in (28) for generation g. To understand the interpretation of this coefficient, consider region A whose broadband internet penetration in 2005 was one standard deviation higher than that of region B. In each year t, β_t is the difference between regions A and B in the relative outcome of cognitive-intensive occupations. Similarly, β_t^c is the differential change in the relative outcome of cognitive-intensive occupations for cohort c.

The consistent estimation of equations (27)–(28) requires an exogenous source of variation on the adoption of broadband internet across German districts in 2005. However, the cross-regional variation in internet penetration is unlikely to be random since adoption should be faster in regions with workers more suitable to use that technology. For instance, this would be the case if broadband internet expands first in regions with a growing number of young individuals specialized in cognitive-intensive occupations. To circumvent this issue, we follow Falck, Gold, and Heblich (2014) to obtain exogenous variation in broadband internet adoption across German districts stemming from pre-existing conditions of the regional telephone networks. In West Germany, the telephone network constructed in the 1960s used copper wires to connect households to the municipality's main distribution frame (MDF). The initial roll-out of DSL internet access in Germany used these pre-existing copper wire lines to provide high-speed internet to households. As argued by Falck, Gold, and Heblich (2014), the copper wire transmission technology did not support high-speed internet provision over long distances. In fact, provision was impossible in areas located more than 4200m away from an existing main distribution frame (MDF). It was necessary to set up an entirely new system to provide DSL access to areas connected to an MDF located more than 4200m away. Thus, areas initially located close to MDFs obtained broadband internet access before areas located far from them.

This discussion suggests that the initial location of MDFs is an exogenous shifter of DSL access in 2005. This requires that, conditional on controls, the determinants of MDF construction in the 1960s were orthogonal to the determinants of changes in labor market outcomes in the 2000s, except through their effect on broadband internet penetration in 2005.[22] Building on this idea, we construct two instrumental variables at the district-level that measure the region's population share located in areas where the existing telephone network could not be used to supply high-speed internet. These variables are aggregates of the municipality-level instrumental variables used in Falck, Gold, and Heblich (2014). The first variable is a simple count of the number of municipalities in the district that did not have a MDF within the municipality, and whose population center (measuring as a population-weighted centroid)

[22] While some of these MDFs were built in population centers, others were built in locations where large empty building sites were available. Falck, Gold, and Heblich (2014) provide a detailed discussion of why the main orthogonality assumption is plausible in this setting. Our strategy is similar to the geographic barriers exploited in Akerman, Gaarder, and Mogstad (2015) to estimate the impact of broadband internet on within-firm skill upgrading in Norway. In contrast, our empirical strategy uncovers reduced-form responses in regional outcomes, which combine adjustment margins within and between firms at the regional-level.

was further than the cut-off threshold of 4200m to the MDF used by the municipality. We refer to this variable as the "MDF density measure." The second variable counts the number of municipalities that satisfied the conditions in the first variable, but were further hampered by the lack of any MDFs in neighboring municipalities that were closer than 4200m. The municipalities in the second group required the installation of completely new networks since it was not possible to install copper wire lines connecting them to *any* existing MDF. We refer to this variable as "Alternative MDF availability."

Let Z_i denote the district-level instrument vector with the district's "MDF density measure" and "Alternative MDF availability". Since the observations in equation (27) vary at the occupation-district level, we estimate this equation with an instrument vector that includes Z_i interacted with a constant and the cognitive-intensity $\bar{C}_o$ for each occupation o. Similarly, to equation (28), we also interact the instrumental variable vector with dummies for each generation g.

5.4.2 Results

We start by examining the first-stage regression that relates the initial telephone network to DSL access. Although equations (27)–(28) vary by district-occupation or district-occupation-generation, the exogenous variation in the instrument vector is only across districts. Therefore, to provide a clear picture of the exogenous variation underlying the model's first-stage we first examine the impact of the instrument vector Z_i on the district's share of population with broadband internet access, DSL_i.[23] That is, we begin by estimating the following linear regression for year t:

$$DSI_{i,t} = Z_i \rho_t + X_i \gamma_t + \epsilon_{i,t} \tag{29}$$

where X_i is the vector of district-level controls used in the estimation of (27)–(28).[24]

Table 2 shows that districts with adverse initial conditions for internet adoption had a lower share of households with high-speed internet in 2005. This difference is smaller by 2007, but it remains significant. Thus, regional differences in broadband penetration are converging throughout the period of analysis. Columns (2) and (4) report the first-stage estimates controlling for the baseline set of district-level controls. We can see that the F statistic of excluded variables remain high, but the alternative MDF availability variable is no longer significant in 2007.

We now turn to the estimation of β_t in (27). Figure 7 reports the estimates of the broadband internet expansion on relative employment (Panel A) and relative payroll (Panel B) in cognitive-intensive occupations. For both outcomes, we find no evidence of responses in the

[23] As discussed above, when estimating equations (27)–(28) we have multiple endogenous variables – DSL access interacted with occupation cognitive intensity and cohort fixed effects. We therefore need to test for weak instruments. We provide the Sanderson-Windmeijer F-statistics (Sanderson and Windmeijer, 2016) for the first stage for each specification in Appendix B. This test statistic checks for whether any of our endogenous variables are weakly instrumented, as well as whether there are sufficiently many strong instruments to instrument the multiple endogenous variables.

[24] Besides these controls, the estimation of (27)–(28) also includes district-occupation-generation pretrends in the dependent variable. In Appendix Table A4, we show that our results are similar in the absence of these controls.

Table 2: First-Stage Regressions – Share of households with DSL access in 2005 and 2007

	2005		2007	
	(1)	(2)	(3)	(4)
MDF density measure	-0.020***	-0.018***	-0.026***	-0.024***
	(0.005)	(0.005)	(0.005)	(0.005)
Alternative MDF availability	0.002	-0.001	0.002*	-0.00003
	(0.001)	(0.002)	(0.002)	(0.001)
Baseline controls	Yes	No	Yes	No
First Stage F	26.49	43.06	23.51	38.62

Notes: Note. Sample of 323 districts in West Germany. All regressions are weighted by the district population size in 1999. Baseline controls include the following district variables in 1999: college graduate population share, manufacturing employment share, immigrant employment share and workforce age composition. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

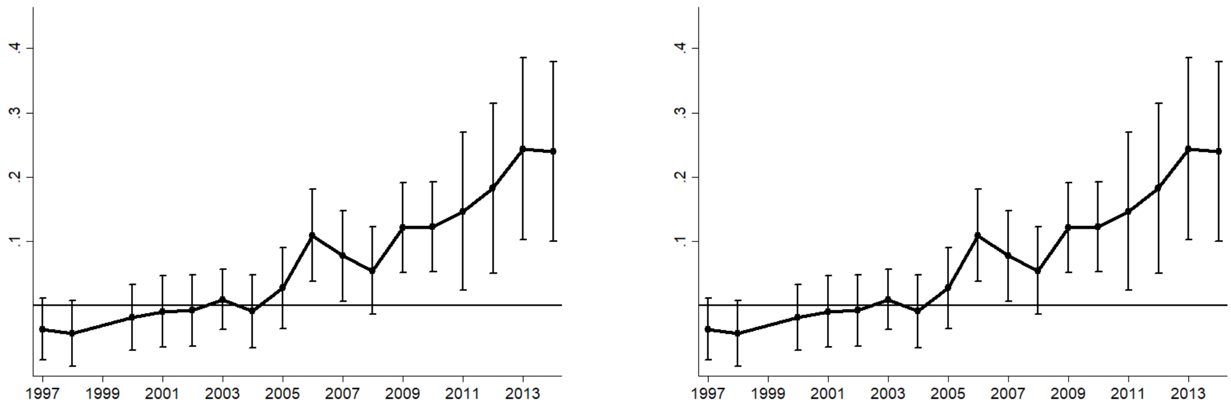
pre-shock period of 1997-2005. Starting in 2005, our estimates indicate a slow and steady increase in the relative employment in cognitive-intensive occupations. In 2014, the point estimate suggests that a region with a one-standard deviation higher broadband internet penetration in 2005 had 0.3 log-points higher employment in the most cognitive-intensive occupation than in the least cognitive-intensive occupation.

These results are consistent with the predictions of our model. We interpret the introduction of broadband internet as a positive shock to the relative productivity of the occupations that use this technology more intensively: cognitive-intensive occupations. In line with the results in Section 3, we find that employment and payroll increase more in these occupations. Such a positive impact becomes larger along the transition to the new stationary equilibrium following the shock.

We now turn to the estimation of employment responses for each generation: β_t^{old} and β_t^{young} obtained from (28). Panel A of Figure 8 reports the estimates for each year between 1997 and 2014. Prior to 2003, regions with early DSL expansion did not experience differential growth in the relative outcomes of cognitive-intensive occupations for old and young workers. After 2005, we find a significant impact on the relative employment of young cohorts in cognitive-intensive occupations. In contrast, we do not find such an effect for old cohorts – if anything, the effect is negative. Panel B of Figure 8 shows that the between-generation difference in relative employment growth is statistically significant in every year after 2006. In line with our model's prediction, the between-generation component grows shortly after the shock and then starts to stabilize.

We can again use our model to interpret the results in Figure 8. The small relative employment response of old generations suggests that the short-run skill supply elasticity is close to zero. In this case, old generations do not switch occupations as their skills would have lower value in the more cognitive intensive occupations augmented by the technological innovation. Alternatively, the positive between-generation difference in employment

Figure 7: Impact of early DSL adoption: All generations

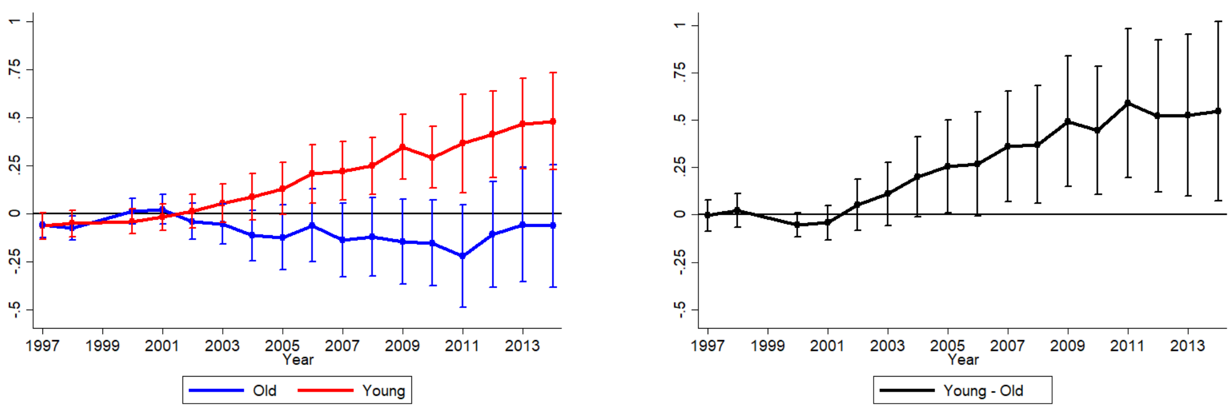


(a) Relative aggregate employment response (b) Relative aggregate wagebill response

Notes: Estimation of equation (27) in the sample of 120 occupations, and 323 districts. Dependent variable: log employment (left) and log payroll (right). All regressions are weighted by the district population size in 1999 and include occupation-time fixed-effects. Baseline controls include the following district variables in 1999: college graduate population share, manufacturing employment share, immigrant employment share, district age composition, and the dependent variable pre-shock growth in 1995-1999. For each year, the dot is the point estimate of β^t, and the bar is associated 90% confidence interval implied by the standard error clustered at the district level.

response indicates that incoming cohorts adapt their skill investment decision towards skills appropriate for performing cognitive-intensive jobs. This suggests that the elasticity of skill supply is positive in the long-run.

Figure 8: Impact of early DSL adoption on Relative Cognitive-intensive Employment: Within- and Between Generations



(a) Relative employment response by generation (b) Between employment response and aggregate

Notes: Estimation of equation (28) in the sample of 2 cohorts, 120 occupations, and 323 districts. Dependent variable: log employment. All regressions are weighted by the district population size in 1999 and include occupation-time and cohort-time fixed-effects. Baseline controls include the following district variables in 1999: college graduate population share, manufacturing employment share, immigrant employment share, district age composition, and the dependent variable pretrend growth in 1995-1999. For each year, the dot is the point estimate of β^t, and the bar is associated 90% confidence interval implied by the robust standard error.

6 Quantitative Analysis

6.1 Calibration

We calibrate our model in two steps. We first specify the functional forms for the innate ability distribution, $\bar{\epsilon}(i)$. Since our model does not have a direct interpretation for the skills i, we select the distribution of innate ability to normalize the initial skill distribution to be uniform: $s_0(i) \equiv 1$.[25] In addition, we abstract from differences across skills in non-cognitive productivity by normalizing $\alpha(i) \equiv 1$. We calibrate the discount rate to match an annual interest rate of 2%, $\rho = 0.02$. Finally, we calibrate the demand elasticity of substitution to $\theta = 3$.

We then calibrate all remaining parameters of the model using the the results from Section 5. We select $\delta = 0.057$ and $\tilde{x}_0 = 40\%$ to match the decline in the share of the old generation in total employment from 40% in 1997 to 15% in 2014. In line with the discussion in Section 4, we select ψ to match the estimated impulse response function of the between-generation difference in relative cognitive-intensive employment. The positive estimated coefficients in Panel B of Figure 8 yields $\psi = 0.35$. In addition, we select η to match the estimated impulse response function of the relative cognitive-intensive employment of the old generation. Due to the nonsignificant estimates reported in Panel B of Figure 8, we calibrate $\eta = 0$. We formally present the calibration procedure in Appendix B.3, along with the in-sample model fit. Finally, for all welfare calculations, we set welfare-weights re^{-rt} with $r = \rho + \delta$, i.e. social discounting of future generations is identical to the discounting of worker's future utility.

6.2 Skill-specificity, welfare, and the economy's adjustment to skill-biased innovations

In our first quantitative exercise, we consider a skill-biased innovation $\Delta log(A)$ that increases the employment share of the H technology from 0.2 in the initial equilibrium to 0.5 in the new long-run equilibrium under our baseline parameterization of an economy like Germany. The first column in Table 3 shows that the consumption equivalent increase in average welfare across generations $\Delta\bar{U} * (\rho + \delta)$ is 46 percent and the lifetime welfare inequality increase is 39 percent in the "Baseline" economy. The "Short-run" row instead shows the changes in average welfare and log-relative wages on impact. The "Long-run" row instead shows these changes across steady states. We find that the short-run calculation severely understates the average welfare gains and overstates the inequality increases. The opposite is true for the long-run calculation, although the magnitudes of the biases are smaller. The biases come precisely from the slow adjustment dynamics that arise in economies with high skill-specificity, as we showed in Section 3.2. This can be seen from the last line of the table which shows the persistence of lifetime inequality $\int_0^\infty \hat{q}_t dt$ is 1.57 average lifetimes (i.e., $1/\delta$).

[25]Under the assumption that the economy is in a stationary equilibrium at $t = 0$, the isoelastic function governing the evolution of the skill lottery in 1 implies that any calibration of the skill distribution does not affect worker choices conditional on wage changes.

In our second quantitative exercise, we consider economies where the skill-specificity is lower but that is subject to the same $\Delta log(A)$ than in our first exercise. In particular, we consider economies where the short-run skill supply elasticity is either 0.5 or 1. The second and third columns in Table 3 show the results. As the specificity of skills becomes smaller, the biases from both the short-run or long-run calculations become smaller as well. The reason for this is that the economy's adjustment dynamics become faster. For example, the persistence goes from 1.57 lifetimes when $\eta = 0$ to 0.71 lifetimes when $\eta = 1$.

Table 3: Welfare changes ($\psi = 0.35$)

	$\eta = 0$		$\eta = 0.5$		$\eta = 1$	
	$\Delta\bar{U} * (\rho + \delta)$	$\Delta\bar{\Omega} * (\rho + \delta)$	$\Delta\bar{U} * (\rho + \delta)$	$\Delta\bar{\Omega} * (\rho + \delta)$	$\Delta\bar{U} * (\rho + \delta)$	$\Delta\bar{\Omega} * (\rho + \delta)$
Baseline	46%	39%	43%	32%	44%	26%
Short-run	31%	76%	38%	53%	42%	39%
Long-run	55%	30%	47%	26%	47%	22%
$\frac{\int_0^\infty \hat{q}_t dt}{1/\delta}$	1.57		1.03		0.71	

7 Conclusion

How do economies adjust to new technologies? We develop an overlapping generations model with worker heterogeneity in technology-specific skills. Forward-looking workers invests in different skills upon entry. Conditional on their skill type, they self-select to work with one technology at each point in time. Following a technology-improving innovation, we characterize in closed-form the dynamic responses of labor market outcomes for different worker generations as a function of parameters governing the short- and long-run skill supply elasticities to changes in relative wages.

We exploit this result in two ways. First, to understand how these parameters shape the distributional welfare consequences of the innovation through its effect on the magnitude and persistence of the adjustment. Second, to measure these parameters by linking them to observable dynamic responses across generations.

We then use our theory to study the adjustment of German regions to the introduction of broadband internet in the early 2000s. We find that, in early adopting regions, the employment in cognitive-intensive occupations increased for young workers but not for old workers. This evidence suggests that the cognitive-skills supply elasticity is low in the short-run and moderate in the long-run. Ignoring the slow adjustment across generations by only considering long-run changes overstates the average welfare benefits and understates the lifetime inequality increase following cognitive-biased technological innovations. Ignoring such slow adjustment by only using observed short-run responses results in the opposite bias.

References

Acemoglu, Daron and David Autor. 2011. "Skills, tasks and technologies: Implications for employment and earnings." In *Handbook of labor economics*, vol. 4. Elsevier, 1043–1171.

Acemoglu, Daron and Pascual Restrepo. 2017. "Robots and Jobs: Evidence from US Labor Markets." Working Paper 23285, National Bureau of Economic Research.

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

Adão, Rodrigo. 2016. "Worker heterogeneity, wage inequality, and international trade: Theory and evidence from brazil." *Unpublished paper, MIT* .

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. 2015. "The Skill Complementarity of Broadband Internet." *The Quarterly Journal of Economics* 130 (4):1781–1824.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New trade models, same old gains?" *American Economic Review* 102 (1):94–130.

Atkeson, Andrew, Ariel Burstein, and Manolis Chatzikonstantinou. 2018. "Transitional Dynamics in Aggregate Models of Innovative Investment." Tech. rep., National Bureau of Economic Research.

Atkeson, Andrew and Patrick J. Kehoe. 2007. "Modeling the Transition to a New Economy: Lessons from Two Technological Revolutions." *American Economic Review* 97 (1):64–88.

Atkin, David. 2016. "Endogenous Skill Acquisition and Export Manufacturing in Mexico." *American Economic Review* 106 (8):2046–85.

Autor, David and David Dorn. 2009. "The Skill Content of Jobs and the Evolution of the Wage Structure-This Job is" Getting Old": Measuring Changes in Job Opportunities using Occupational Age Structure." *American Economic Review* 99 (2):45.

Autor, David and Anna Salomons. 2017. "Robocalypse now: does productivity growth threaten employment?" *NBER Chapters* .

Autor, David H. and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5):1553–97.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics* 118 (4):1279–1333.

Ben-Porath, Yoram. 1967. "The production of human capital and the life cycle of earnings." *Journal of political economy* 75 (4, Part 1):352–365.

Blanchard, Olivier J. 1985. "Debt, deficits, and finite horizons." *Journal of political economy* 93 (2):223–247.

Buera, Francisco J and Joseph P Kaboski. 2012. "The rise of the service economy." *American Economic Review* 102 (6):2540–69.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel. 2016. "Changes in between-group inequality: computers, occupations, and international trade." Tech. rep., mimeo.

Card, David, Joerg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *The Quarterly Journal of Economics* 128 (3):967–1015.

Card, David and Thomas 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.

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil. 2011. "Estimating marginal returns to education." *American Economic Review* 101 (6):2754–81.

Caselli, Francesco. 1999. "Technological revolutions." *American economic review* 89 (1):78–102.

Chari, Varadarajan V and Hugo Hopenhayn. 1991. "Vintage human capital, growth, and the diffusion of new technology." *Journal of political Economy* 99 (6):1142–1165.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo. Forthcoming. "Housing Booms and Busts, Labor Market Opportunities, and College Attendance." *American Economic Review* .

Costinot, Arnaud and Jonathan Vogel. 2010. "Matching and inequality in the world economy." *Journal of Political Economy* 118 (4):747–786.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum. 2014. "The rise of the East and the Far East: German labor markets and trade integration." *Journal of the European Economic Association* 12 (6):1643–1675.

Dix-Carneiro, Rafael and Brian K. Kovak. 2017. "Trade Liberalization and Regional Dynamics." *American Economic Review* 107 (10):2908–46.

Durlauf, Steven N and Philippe Aghion. 2005. "Handbook of economic growth." .

Eaton, Jonathan and Samuel Kortum. 2002. "Technology, geography, and trade." *Econometrica* 70 (5):1741–1779.

Falck, Oliver, Robert Gold, and Stephan Heblich. 2014. "E-lections: Voting Behavior and the Internet." *American Economic Review* 104 (7):2238–65.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M Shapiro. 2018. "Pre-event Trends in the Panel Event-study Design." Working Paper 24565, National Bureau of Economic Research.

Galor, Oded and Omer Moav. 2002. "Natural Selection and the Origin of Economic Growth." *The Quarterly Journal of Economics* 117 (4):1133–1191.

Hansen, Lars Peter and Thomas J Sargent. 2008. *Robustness*. Princeton university press.

Hayashi, Fumio. 1982. "Tobin's marginal q and average q: A neoclassical interpretation." *Econometrica: Journal of the Econometric Society* :213–224.

Herrendorf, Berthold, Christopher Herrington, and Akos Valentinyi. 2015. "Sectoral technology and structural transformation." *American Economic Journal: Macroeconomics* 7 (4):104–33.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi. 2014. "Growth and structural transformation." In *Handbook of economic growth*, vol. 2. Elsevier, 855–941.

Hobijn, Bart, Todd Schoellman, and Alberto Vindas. 2019. "Wages during Structural Transformation: The Importance of Cohort Labor Supply Decisions." .

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow. 2013. "The allocation of talent and us economic growth." Tech. rep., National Bureau of Economic Research.

Huber, Kilian. 2018. "Disentangling the Effects of a Banking Crisis: Evidence from German Firms and Counties." *American Economic Review* 108 (3):868–98. URL http://www.aeaweb.org/articles?id=10.1257/aer.20161534.

Katz, Lawrence F and Kevin M Murphy. 1992. "Changes in relative wages, 1963–1987: supply and demand factors." *The quarterly journal of economics* 107 (1):35–78.

Kim, Dae-Il and Robert H Topel. 1995. "Labor markets and economic growth: lessons from Korea's industrialization, 1970-1990." In *Differences and changes in wage structures*. University of Chicago Press, 227–264.

Lagakos, David and Michael E. Waugh. 2013. "Selection, Agriculture, and Cross-Country Productivity Differences." *American Economic Review* 103 (2):948–80.

Lee, Donghoon and Kenneth I Wolpin. 2006. "Intersectoral labor mobility and the growth of the service sector." *Econometrica* 74 (1):1–46.

Matsuyama, Kiminori. 1992. "A simple model of sectoral adjustment." *The Review of Economic Studies* 59 (2):375–387.

Moll, Benjamin, Lukasz Rachel, and Pascual Restrepo. 2019. "Uneven Growth: Automation's Impact on Income and Wealth Inequality." .

Nakamura, Emi and Jón Steinsson. 2017. "Identification in macroeconomics." Tech. rep., National Bureau of Economic Research.

Ngai, L Rachel and Christopher A Pissarides. 2007. "Structural change in a multisector model of growth." *American economic review* 97 (1):429–443.

Porzio, Tommaso and Gabriella Santangelo. 2019. "Does Schooling Cause Structural Transformation?" .

Roy, Andrew Donald. 1951. "Some thoughts on the distribution of earnings." *Oxford economic papers* 3 (2):135–146.

Sanderson, Eleanor and Frank Windmeijer. 2016. "A weak instrument F-test in linear IV models with multiple endogenous variables." *Journal of Econometrics* 190 (2):212–221.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of monetary Economics* 50 (3):665–690.

Spitz-Oener, Alexandra. 2006. “Technical change, job tasks, and rising educational demands: Looking outside the wage structure.” *Journal of labor economics* 24 (2):235–270.

Tobin, James. 1969. “A general equilibrium approach to monetary theory.” *Journal of money, credit and banking* 1 (1):15–29.

Yaari, Menahem E. 1965. “Uncertain lifetime, life insurance, and the theory of the consumer.” *The Review of Economic Studies* 32 (2):137–150.

Young, Alwyn. 2014. “Structural transformation, the mismeasurement of productivity growth, and the cost disease of services.” *American Economic Review* 104 (11):3635–67.

Appendix A Proofs

A.1 Proof of Lemma 2

We obtain (14) by applying this expression into the relative supply expression in (13) and the relative demand expression in (2). Existence and uniqueness follow from applying Bolzano's theorem to (14). The left-hand side is strictly decreasing in ω_t, converges to zero as $\omega_t \to \infty$, and converges to infinity as $\omega_t \to 0$. Notice that $l(\omega_t)$ is decreasing in ω_t and $l(\omega_t) \in [0,1]$. Thus, the right-hand-side is strictly increasing in $l(\omega_t)$, it converges to infinity as $l(\omega_t) \to 0$ and it converges to zero if $l(\omega_t) \to 1$.

A.2 Proof Lemma 3

The FOC of workers' skill-accumulation problem are:

$$V_t(i) - \frac{1}{\psi}\left(1 + log\left(\frac{\tilde{s}_t(i)}{\bar{s}_t(i)}\right)\right) - \lambda_t = 0$$

$$\lambda_t \left(\int_0^1 \tilde{s}_t(x)dx - 1\right) = 0$$

Integrating over $i \in [0,1]$, we obtain an equation characterizing λ_t:

$$log\left(\int_0^1 \bar{s}_t(i)e^{\psi V_t(i)}di\right) = \psi\lambda_t + 1$$

Therefore,

$$\tilde{s}_t(i) = \frac{\bar{s}_t(i)e^{\psi V_t(i)}}{\int_0^1 \bar{s}_t(j)e^{\psi V_t(j)}dj}.$$

Using the wage expressions and assignment function in Lemma 1, we can write the value function of a worker i at time t as

$$\begin{aligned}
V_t(i) &= \int_t^\infty e^{-(\rho+\delta)(s-t)} log(w_s(i))ds \\
&= \int_t^\infty e^{-(\rho+\delta)(s-t)} \left(log(\omega_s\sigma(i)\alpha(i))\mathbb{I}_{i\geq l_s} + log(\alpha(i))\left(1-\mathbb{I}_{i<l_s}\right)\right) ds \\
&= \frac{log(\alpha(i))}{\rho+\delta} + \int_t^\infty e^{-(\rho+\delta)(s-t)} log\left(\omega_s\sigma(i)\right)\mathbb{I}_{i\geq l_s} ds
\end{aligned}$$

By defining $Q_t(i) \equiv e^{\int_t^\infty e^{-(\rho+\delta)(s-t)} log(\omega_s\sigma(i))\mathbb{I}_{i\geq l_s} ds}$, we obtain

$$\tilde{s}_t(i) = \frac{\bar{s}_t(i)\alpha(i)^{\frac{\psi}{\rho+\delta}}Q_t(i)^\psi}{\int_0^1 \bar{s}_t(j)\alpha(j)^{\frac{\psi}{\rho+\delta}}Q_t(j)^\psi dj}.$$

A.3 Proof of Proposition 1

First, we do a first order approximation around the stationary equilibrium of equations (10), (12) and (14). We obtain:

$$\frac{\partial \hat{s}_t(i)}{\partial t} = -\delta \hat{s}_t(i) + \delta \hat{\tilde{s}}_t(i) \tag{A.1}$$

$$\hat{l}_t = \frac{\eta}{\theta - 1}\hat{y}_t \tag{A.2}$$

$$\hat{l}_t = \frac{\eta}{\kappa\eta + \theta}\left(\int_l^1 \hat{s}_t(i)\frac{\alpha(i)\sigma(i)s(i)}{\int_l^1 \alpha(i)\sigma(i)s(i)di}di - \int_0^l \hat{s}_t(i)\frac{\alpha(i)s(i)}{\int_0^l \alpha(i)s(i)di}di\right) \tag{A.3}$$

where

$$\kappa \equiv \frac{\alpha(l)s(l)l}{\int_0^l \alpha(i)s(i)di} + \frac{\alpha(l)\sigma(l)s(l)l}{\int_l^1 \alpha(i)\sigma(i)s(i)di}.$$

Differentiating (A.3) annd (A.2) with respect to time, we get that

$$\frac{\partial \hat{y}_t}{\partial t} = \frac{\theta - 1}{\kappa\eta + \theta}\left(\int_l^1 \frac{\partial \hat{s}_t(i)}{\partial t}\frac{\alpha(i)\sigma(i)s(i)}{\int_l^1 \alpha(i)\sigma(i)s(i)di}di - \int_0^l \frac{\partial \hat{s}_t(i)}{\partial t}\frac{\alpha(i)s(i)}{\int_0^l \alpha(i)s(i)di}di\right)$$

Applying (A.1) to this expression, we obtain

$$\frac{\partial \hat{y}_t}{\partial t} = -\delta\hat{y}_t + \frac{\theta - 1}{\kappa\eta + \theta}\delta\left(\int_l^1 \hat{\tilde{s}}_t(i)\frac{\alpha(i)\sigma(i)s(i)}{\int_l^1 \alpha(i)\sigma(i)s(i)di}di - \int_0^l \hat{\tilde{s}}_t(i)\frac{\alpha(i)s(i)}{\int_0^l \alpha(i)s(i)di}di\right). \tag{A.4}$$

Furthermore, we will guess an verify that l_t converges monotonically along the equilibrium path. We show the proof starting from $\hat{l}_0 < 0$. The proof for $\hat{l}_0 > 0$ is analogous and omitted.

Whenever $\hat{l}_0 < 0$ and increases monotonically along the equilibrium path, we have that for all $s > t$, types $i < l_t$ are employed in technology L and types $i > l$ are employed in technology H. Also, for workers with $i \in (l_t, l)$, there exist a $\tau(i)$ such that they work in H for all $t < s < t + \tau(i)$ and in L for all $s > t + \tau(i)$.

Then, from equation (15), we have

$$Q_t(i) = \begin{cases} 1 & i \leq l_t \\ e^{\int_t^{t+\tau(i)} e^{-(\rho+\delta)(s-t)} log(\omega_s \sigma(i))ds} & i \in (l_t, l) \\ \sigma(i)^{\frac{1}{\rho+\delta}} q_t & i > l \end{cases} \tag{A.5}$$

So, we can write the optimal lottery as

$$\tilde{s}_t(i) = \begin{cases} \frac{\tilde{s}(i)}{\tilde{s}(l)}\tilde{s}_t(l)e^{-\psi\int_t^\infty e^{-(\rho+\delta)(s-t)} log\left(\frac{\omega_s}{\omega}\right)ds} & i \leq l_t \\ \frac{\tilde{s}(i)}{\tilde{s}(l)}\left(\frac{\sigma(i)}{\sigma(l)}\right)^{\frac{\psi}{\rho+\delta}\left(1-e^{-(\rho+\delta)\tau(i)}\right)} \tilde{s}_t(l)e^{-\psi\int_{t+\tau(i)}^\infty e^{-(\rho+\delta)(s-t)} log\left(\frac{\omega_s}{\omega}\right)ds} & i \in (l_t, l) \\ \frac{\tilde{s}(i)}{\tilde{s}(l)}\tilde{s}_t(l) & i \geq l \end{cases} \tag{A.6}$$

Log-linearizing (A.6) we obtain that

$$\hat{\hat{s}}_t(i) = \hat{\hat{s}}_t(l) - \psi\hat{q}_t\mathbb{I}_{i<l_t} - \psi\hat{q}_{t+\tau(i)}\mathbb{I}_{i\in(l_t,l)} \tag{A.7}$$

Replacing in the expression inside the parenthesis in (A.4), we obtain

$$\begin{aligned}
\left(\int_l^1 \hat{\hat{s}}_t(i)\frac{\alpha(i)\sigma(i)s(i)}{\int_l^1 \alpha(i)\sigma(i)s(i)di}di - \int_0^l \hat{\hat{s}}_t(i)\frac{\alpha(i)s(i)}{\int_0^l \alpha(i)s(i)di}di\right) &= \\
\int_0^l \psi\left(\hat{q}_t\mathbb{I}_{i<l_t} + \hat{q}_{t+\tau(i)}\mathbb{I}_{i>l_t}\right)\frac{\alpha(i)s(i)}{\int_0^l \alpha(x)s(x)dx}di &= \\
\psi\hat{q}_t - \psi\int_{l_t}^l\left(\hat{q}_t - \hat{q}_{t+\tau(i)}\right)\frac{\alpha(i)s(i)}{\int_0^l \alpha(x)s(x)dx}di &
\end{aligned}$$

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

Then, given our guess that l_t increases monotonically along the equilibrium path, from (12) we see that ω_t decreases monotonically along the equilibrium path. This implies that $\hat{q}_t > \hat{q}_{t+\tau(i)} > 0$ for all i and all t. So, we can show that the term inside the integral is of second order:

$$0 \leq \int_{l_t}^l\left(\hat{q}_t - \hat{q}_{t+\tau(i)}\right)\frac{\alpha(i)s(i)}{\int_0^l \alpha(x)s(x)dx}di \leq \int_{l_t}^l \hat{q}_t\frac{\alpha(i)s(i)}{\int_0^l \alpha(x)s(x)dx}di \leq \frac{\max_{i\in(l_t,l)}\alpha(i)s(i)l}{\int_0^l \alpha(x)s(x)dx}\hat{l}_t\hat{q}_t \approx 0.$$

Replacing this expression back in (A.4), we obtain the Kolmogorov-Forward equation for $\hat{y}_t$ shown in the lemma,

$$\frac{\partial\hat{y}_t}{\partial t} = -\delta\hat{y}_t + \frac{\theta-1}{\kappa\eta+\theta}\delta\psi\hat{q}_t. \tag{A.8}$$

To show the Kolmogorov-Backward equation satisfied by $\hat{q}_t$, we differentiate with respect to time and obtain

$$\frac{\partial q_t}{\partial t} = -\omega_t + (\rho+\delta)q_t.$$

Then, using that $p_t = (y_t)^{\frac{1}{1-\theta}}$ and log-linearizing, we obtain the equation shown in the lemma,

$$\frac{\partial\hat{q}_t}{\partial t} = \frac{1}{\theta-1}\hat{y}_t + (\rho+\delta)\hat{q}_t. \tag{A.9}$$

To complete the proof, we need to derive the policy functions, show the equilibrium is saddle-path stable, and verify that l_t increases monotonically along the equilibrium path (which we guessed in order to derive the 2X2 dynamic system in $\hat{q}_t, \hat{y}_t$).

Let us guess that the policy functions are given by $\frac{\partial\hat{y}_t}{\partial t} = -\lambda\hat{y}_t$ and $\hat{q}_t = \zeta\hat{y}_t$. Replacing in the 2X2 dynamic system, we obtain the expressions in the proposition for λ and ζ:

$$
\begin{aligned}
-\lambda =& -\delta + \frac{\theta-1}{\kappa\eta+\theta}\delta\psi\zeta \\
-\zeta\lambda =& \frac{1}{\theta-1} + (\rho+\delta)\zeta
\end{aligned}
$$

Notice that the second equation immediately yield the expression for ζ. To get the expression for λ, notice that substituting the expression for ζ into the first equation implies that

$$
(\delta-\lambda)(\rho+\delta+\lambda) + \frac{\psi\delta}{\kappa\eta+\theta} = 0
$$

Then, the solutions for λ is

$$
\lambda_{12} = -\frac{\rho}{2} \pm \sqrt{\left(\frac{\rho}{2}\right)^2 + \delta\left((\rho+\delta) + \frac{\psi}{\kappa\eta+\theta}\right)}
$$

Because the term inside the square root is always positive, two solutions always exist. Furthermore, one of the solutions is negative and the other one is positive. This implies that the equilibrium is saddle-path stable. Furthermore, the positive solution is the speed of convergence of equilibrium variables.

Finally, the equilibrium threshold is $\hat{l}_t = \hat{l}_0 e^{-\lambda t}$. Then, if $\hat{l}_0 < 0$, this implies that l_t increases monotonically along the equilibrium path, which verifies our initial guess and completes the proof of the proposition.

A.4 Proof of Corollary 1

Notice that $\int s(i)\hat{\tilde{s}}_t(i)di = \int(\tilde{s}_t(i) - s(i))di = 0$. Using (A.7) , we have that

$$
\begin{aligned}
0 =& \int_0^1 s(i)\hat{\tilde{s}}_t(i)di \\
=& \hat{\tilde{s}}_t(l) - \psi\int_0^l \left(\hat{q}_t\mathbb{I}_{i<l_t} + \hat{q}_{t+\tau(i)}\mathbb{I}_{i\in(l_t,l)}\right)s(i)di \\
=& \hat{\tilde{s}}_t(l) - \left(\int_0^l s(i)di\right)\psi\hat{q}_t + \psi\int_{l_t}^l \left(\hat{q}_t - \hat{q}_{t+\tau(i)}\right)s(i)di
\end{aligned}
$$

We can use use the same arguments as in A.3 to show that the last term is of second order. Thus,

$$
\hat{\tilde{s}}_t(l) = \left(\int_0^l s(i)di\right)\psi\hat{q}_t
$$

and, therefore,

$$
\hat{\tilde{s}}_t(i) = \left(\int_0^l s(i)di\right)\psi\hat{q}_t - \psi\hat{q}_t\mathbb{I}_{i<l} + \psi(\hat{q}_t - \hat{q}_{t+\tau(i)})\mathbb{I}_{i\in(l_t,l)}.
$$

To prove the result, we use that $\hat{q}_{t+\tau(i)} = \hat{q}_t e^{-\lambda\tau(i)}$. So,

$$
\begin{aligned}
\hat{\hat{s}}_t(i) &= \left(\int_0^l s(i)di\right)\psi\hat{q}_t - \psi\hat{q}_t\mathbb{I}_{i<l} + \psi(\hat{q}_t - \hat{q}_{t+\tau(i)})\mathbb{I}_{i\in(l_t,l)} \\
&= \mathbb{I}_{i>l}\psi\hat{q}_t - \left(1 - \int_0^l s(i)di\right)\psi\hat{q}_t + \psi\hat{q}_t(1-e^{-\lambda\tau(i)})\mathbb{I}_{i\in(l_t,l)} \\
&= \left(\mathbb{I}_{i>l} - \int_l^1 s(i)di\right)\psi\hat{q}_t + o_t(i)
\end{aligned}
$$

where $o_t(i) \equiv \psi\hat{q}_t(1-e^{-\lambda\tau(i)})\mathbb{I}_{i\in(l_t,l)}$ and has $\int s(i)o_t(i)di = 0$.

Finally, the dynamics of the skill-distribution and the threshold were already derived in equations A.1 and A.2.

A.5 Proof of Proposition 2

Using the definitions y_t and q_t together with Proposition 1, we have

$$
\begin{aligned}
\Delta log(y_t) &= (\theta-1)\left(\Delta log(A) - \Delta log(\omega) - \hat{\omega}_t\right) \\
&= (\theta-1)\left(\Delta log(A) - \left(\Delta log(\omega) + \hat{\omega}_0 e^{-\lambda t}\right)\right) \\
\Delta log(q_t) &= \Delta log(q) + \hat{q}_t \\
&= \frac{1}{\rho+\delta}\Delta log(\omega) + \frac{1}{\rho+\delta+\lambda}\hat{\omega}_0 e^{-\lambda t}
\end{aligned}
$$

We next derive the long-run change $\Delta log(\omega)$ and the short-to-long-run change $\hat{\omega}_0$

Long-run. In this case the skill distribution is given by (16), so that the equilibrium threshold solves

$$
A^{\theta-1}\sigma(l)^\theta \int_0^l \alpha(i)(\alpha(i))^{\frac{\psi}{\rho+\delta}}di = \int_l^1 \alpha(i)\sigma(i)\left(\alpha(i)\frac{\sigma(i)}{\sigma(l)}\right)^{\frac{\psi}{\rho+\delta}}di
$$

Consider a log-linear approximation around the final stationary equilibrium:

$$
(\theta-1)\Delta log(A) + \left(\left(\theta + \frac{\psi}{\rho+\delta}\right)\frac{1}{\eta} + \kappa\right)\Delta log(l) = 0
$$

Thus,

$$
\Delta log(l) = -\frac{\eta}{\left(\theta + \frac{\psi}{\rho+\delta}\right) + \eta\kappa}\left(\theta-1\right)\Delta log(A)
$$

From equation (12), $\Delta log(\omega) = -\frac{1}{\eta}\Delta log(l)$ and, therefore,

$$
\Delta log(\omega) = \frac{1}{\left(\theta + \frac{\psi}{\rho+\delta}\right) + \eta\kappa}\left(\theta-1\right)\Delta log(A) \tag{A.10}
$$

Short-to-Long We start by deriving the change in the skill distribution using (16): $\hat{s}_0(i) = \hat{s}_0(l)$ if $i < l$ and $\hat{s}_0(i) = \hat{s}_0(l) - \frac{\psi}{\rho+\delta}\Delta log(\omega)$ if $i > l$. Along the transition, the change in the assignment threshold is determined by (14) given the change in the skill distribution:

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

Then,

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

Dynamic responses We now use the derivations above to show that

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

where the last line uses the solution to λ from Proposition 1.

A.6 Proof of Proposition 3 and Proposition 4

1. Long-run adjustment

$$\Delta log(y_\infty) = \left(1 - \frac{\theta-1}{\theta+\kappa\eta+\frac{\psi}{\rho+\delta}}\right)(\theta-1)\Delta log(A)$$
$$\Delta log\,(q_\infty) = \frac{1}{\rho+\delta}\left(\frac{1}{1-\theta}\Delta log(y_\infty) + \Delta log(A)\right)$$

Then, it is straightforward to see that $\Delta log(y_\infty)$ is increasing in both η and ψ, while the opposite holds for $\Delta log(q_\infty)$.

2. Short-run adjustment

$$\Delta log(y_0) = \left(1 - \frac{\theta-1}{\theta+\kappa\eta}\right)(\theta-1)\Delta log(A)$$
$$\Delta log\,(q_0) = \frac{1}{\theta+\kappa\eta+\frac{\psi}{\rho+\delta}}\frac{\lambda}{\delta}\frac{1}{\rho+\delta}(\theta-1)\Delta log(A)$$
$$= \frac{1}{\theta+\kappa\eta}\frac{1}{\rho+\lambda}(\theta-1)\Delta log(A)$$

The first line shows that $\Delta log(y_0)$ is increasing η and is independent of ψ. Since λ

is decreasing in η, the second line shows that $\Delta log(q_0)$ is decreasing in η. Since λ is increasing in ψ, the third line shows that $\Delta log(q_0)$ is decreasing in ψ.

3. Persistence

$$\int_0^{\infty} |\hat{y}_t| dt = -\frac{1}{\lambda}\hat{y}_0 = \frac{1}{\lambda}\frac{\frac{\psi}{\rho+\delta}}{\theta+\kappa\eta+\frac{\psi}{\rho+\delta}}\frac{\theta-1}{\theta+\kappa\eta}\left(\theta-1\right)\Delta log(A)$$
$$\int_0^{\infty} \hat{q}_t dt = \frac{1}{\lambda}\hat{q}_0 = \frac{1}{\theta+\eta\kappa+\frac{\psi}{\rho+\delta}}\frac{\lambda-\delta}{\lambda}\frac{1}{\delta}\frac{1}{\rho+\delta}\left(\theta-1\right)\Delta log(A)$$

The second line shows that $\int_0^{\infty}\hat{q}_t dt$ is increasing in ψ around $\psi = 0$ (since then $\lambda = \delta$) and is decreasing in η because λ is decreasing in η and increasing in ψ.

The first line shows that $\int_0^{\infty}|\hat{y}_t|dt$ is increasing in ψ around $\psi = 0$ since $\frac{\partial\left(\frac{1}{\lambda}\frac{\frac{1}{\rho+\delta}}{\theta+\kappa\eta+\frac{\psi}{\rho+\delta}}\right)}{\partial\psi}$ is bounded. To show that it is decreasing in η, we show that:

$$\begin{aligned}\frac{\partial log(\frac{1}{\lambda}\frac{\frac{\psi}{\rho+\delta}}{\theta+\kappa\eta+\frac{\psi}{\rho+\delta}}\frac{\theta-1}{\theta+\kappa\eta})}{\partial\eta} &= \frac{1}{\lambda}\frac{1}{\rho+2\lambda}\frac{\psi\delta\kappa}{(\theta+\kappa\eta)^2} - \frac{\kappa}{\theta+\kappa\eta+\frac{\psi}{\rho+\delta}} - \frac{\kappa}{\theta+\kappa\eta} \\ &= -\left(\left(1-\underbrace{\frac{\lambda-\delta}{\lambda}\frac{\rho+\delta+\lambda}{\rho+2\lambda}}_{<1 \text{ because } \lambda>\delta}\right)\frac{1}{(\theta+\kappa\eta)} + \frac{1}{\theta+\kappa\eta+\frac{\psi}{\rho+\delta}}\right)\kappa < 0\end{aligned}$$

A.7 Proof of Proposition 5

We have that, because of the envelope theorem,

$$\begin{aligned}U_\tau &= \int \tilde{s}_\tau(i)V_t(i)di - \frac{1}{\psi}\int \tilde{s}_\tau(i) log\left(\frac{\tilde{s}_\tau(i)}{\bar{s}(i)}\right)di \\ &\approx \int s(i)(V_\tau(i) - V(i))di + U_\infty\end{aligned}$$

Then,

$$\begin{aligned}U_\tau - U_\infty &= \int_\tau^{\infty} e^{-(\rho+\delta)(t-\tau)}\int s(i) log\left(\frac{\alpha(i)\max(\omega_t\sigma(i),1)}{P_t}\right)didt - \int_0^{\infty} e^{-(\rho+\delta)t}\int s(i) log\left(\frac{\alpha(i)\max(\omega\sigma(}{P}\right. \\ &\approx \int_l^1 s(i)di\left(\int_\tau^{\infty} e^{-(\rho+\delta+\lambda)(t-\tau)}\hat{\omega}_\tau dt\right) - \left(\int_\tau^{\infty} e^{-(\rho+\delta+\lambda)(t-\tau)}\hat{P}_\tau dt\right) \\ &= -\left(\frac{y_\infty}{1+y_\infty}\frac{1}{1-\theta}\hat{y}_0 - \int_l^1 s(i)di\hat{\omega}_0\right)\frac{1}{\rho+\delta+\lambda}e^{-\lambda\tau} \\ &= -\left(\frac{y_\infty}{1+y_\infty} - \int_l^1 s(i)di\right)\hat{q}_\tau\end{aligned}$$

Also,

$$
\begin{aligned}
U_\infty - U_{0^-} &\approx \left(\int_l^1 s(i)di\right)\frac{1}{\rho+\delta}\Delta log\left(\omega_\infty\right) + \frac{y_\infty}{1+y_\infty}\frac{1}{\theta-1}\frac{1}{\rho+\delta}\Delta log(y_\infty) \\
&= \left(\int_l^1 s(i)di\right)\frac{1}{\rho+\delta}\Delta log\left(\omega_\infty\right) + \frac{y_\infty}{1+y_\infty}\frac{1}{\rho+\delta}\left(\Delta log(A) - \Delta log(\omega_\infty)\right) \\
&= \frac{y_\infty}{1+y_\infty}\frac{1}{\rho+\delta}\Delta log(A) - \left(\frac{y_\infty}{1+y_\infty} - \int_l^1 s(i)di\right)\Delta log\left(q_\infty\right)
\end{aligned}
$$

Then,

$$
\begin{aligned}
\Delta\bar{U} &= U_\infty - U_{0^-} + r\int_0^\infty e^{-rt}(U_\tau - U_\infty)d\tau \\
&\approx U_\infty - U_{0^-} - \left(\frac{y_\infty}{1+y_\infty} - \int_l^1 s(i)di\right) r\int_0^\infty e^{-rt}\hat{q}_\tau \\
&= \frac{y_\infty}{1+y_\infty}\frac{1}{\rho+\delta}\Delta log(A) - \left(\frac{y_\infty}{1+y_\infty} - \int_l^1 s(i)di\right)\Delta\bar{\Omega}
\end{aligned}
$$

Finally, using Proposition 2,

$$
\begin{aligned}
\Delta\bar{\Omega} &= r\int_0^\infty e^{-rt}\Delta log(q_\tau)d\tau \\
&= \Delta log(q_\infty) + r\int_0^\infty e^{-rt}\hat{q}_\tau d\tau \\
&\approx \Delta log(q_\infty) + \frac{r}{r+\lambda}\hat{q}_0 \\
&\approx \Delta log(q_\infty) + \frac{r\lambda}{r+\lambda}\int_0^\infty \hat{q}_\tau d\tau
\end{aligned}
$$

A.8 Proof of equations (24)–(25)

Proof of equation (24). We first use a first-order approximation to write the log-change in relative high-tech employment in terms of changes in the high-tech employment share:

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

where $e_{H,t}^{old} = \int_{l_t}^1 s_0(i)di$.

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

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

Notice that

$$
e_{H,t}^{old} - e_{H,0^-}^{old} = \int_l^{l_{0^-}} s_0(i)di + \int_{l_t}^l s_0(i)di
$$

By approximating these expressions around l,

$$\begin{aligned} e^{old}_{H,t} - e^{old}_{H,0^-} &\approx s_0(l)l\left(\Delta\log(l) - \hat{l}_t\right) \\ &\approx (s_0(l)l)\,\eta\Delta\log(\omega_t) \\ &\approx (s_0(l_{0^-})l_{0^-})\,\eta\Delta\log(\omega_t) \\ &\approx (1-e_{H,0^-})\eta\Delta\log(\omega_t) \end{aligned}$$

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

Combining the two expressions,

$$\Delta\log\left(e^{old}_t\right) \approx \frac{1}{e_{H,0^-}}\eta\Delta\log(\omega_t)$$

Using the demand expression in (2),

$$\Delta\log\left(e^{old}_t\right) \approx \frac{1}{e_{H,0^-}}\eta\left(-\frac{1}{\theta-1}\log y_t + \Delta\log A\right)$$

Using the expression for the evolution of y_t in Proposition 2,

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

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

which is identical to (24).

Proof of equation (25). We first use a first-order approximation to write the log-change in relative high-tech employment in terms of changes in the high-tech employment share:

$$\begin{aligned} \log\left(\frac{e^{young}_t}{e^{young}_{0^-}}\right) - \log\left(\frac{e^{old}_t}{e^{old}_{0^-}}\right) \approx& \frac{1}{1-e_{H,\infty}}\left(\frac{e^{young}_{H,t} - e^{young}_{H,0^-}}{e_{H,\infty}} - \frac{e^{old}_{H,t} - e^{old}_{H,0^-}}{e_{H,\infty}}\right) \\ =& \frac{1}{(1-e_{H,\infty})e_{H,\infty}}\left(\left(e^{young}_{H,t} - e^{old}_{H,t}\right) - \left(e^{young}_{H,0^-} - e^{old}_{H,0^-}\right)\right) \\ = \approx& \frac{1}{(1-e_{H,\infty})e_{H,\infty}}\left(e^{young}_{H,t} - e^{old}_{H,t}\right) \end{aligned}$$

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

Using the definition of employment shares for each generation,

$$e_{H,t}^{young} - e_{H,t}^{old} \approx \frac{1}{1-(1-\tilde{x}_0)e^{-\delta t}} \left(\tilde{x}_0 e^{-\delta t} \int_{l_t}^{1} s_0(i)di + \delta \int_0^t e^{\delta(\tau-t)} \int_{l_t}^{1} \tilde{s}_\tau(i) did\tau \right) - \int_{l_t}^{1} s_0(i)di$$
$$\approx \frac{1}{1-(1-\tilde{x}_0)e^{-\delta t}} \left(\delta \int_0^t e^{\delta(\tau-t)} \int_{l_t}^{1} \left(\tilde{s}_\tau(i) - s_0(i)\right) did\tau \right)$$

Thus,

$$\log\left(\frac{e_t^{young}}{e_{0^-}^{young}}\right) - \log\left(\frac{e_t^{old}}{e_{0^-}^{old}}\right) \approx \frac{1}{(1-e_{H,\infty})e_{H,\infty}} \frac{1}{1-(1-\tilde{x}_0)e^{-\delta t}} \left(\delta \int_0^t e^{\delta(\tau-t)} \int_{l_t}^{1} \left(\tilde{s}_\tau(i) - s_0(i)\right) did\tau \right) \tag{A.11}$$

We now consider the following approximation:

$$\int_{l_t}^{1} \tilde{(}\tilde{s}_\tau(i) - s_0(i))di \approx \int_{l}^{1} s(i)(\hat{\tilde{s}}_\tau(i) - \hat{s}_0(i))di$$

Then, we derive $\hat{s}_0(i)$ using the expression for the stationary skill distribution

$$s_0(i) = \frac{\bar{s}(i)\alpha(i)^{\frac{\psi}{\rho+\delta}} \left(\omega_{0^-}\sigma(i)\right)^{\frac{\psi}{\rho+\delta}\mathbb{I}_{i>l_{0^-}}}}{\int_0^{l_{0^-}} \bar{s}(j)\alpha(j)^{\frac{\psi}{\rho+\delta}} dj + \int_{l_{0^-}}^{1} \bar{s}(j)\alpha(j)^{\frac{\psi}{\rho+\delta}} \left(\omega_{0^-}\sigma(j)\right)^{\frac{\psi}{\rho+\delta}} dj}$$
$$\Longrightarrow$$
$$\hat{s}_0(i) \approx -\left(\mathbb{I}_{i>l} - \int_l^1 s(j)dj\right) \frac{\psi}{\rho+\delta} \Delta\log(\omega)$$

Using Corollary 1,

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

We now apply this expression into (A.11):

$$\log\left(\frac{e_t^{young}}{e_{0^-}^{young}}\right) - \log\left(\frac{e_t^{old}}{e_{0^-}^{old}}\right) \approx \frac{\psi}{1-(1-\tilde{x}_0)e^{-\delta t}} \left(\delta \int_0^t e^{\delta(\tau-t)} \left(\hat{q}_\tau + \Delta\log(q)\right) d\tau \right)$$
$$\approx \frac{\psi}{1-(1-\tilde{x}_0)e^{-\delta t}} \left(\delta \int_0^t e^{\delta(\tau-t)} \hat{q}_0 e^{-\lambda\tau} d\tau + (1-e^{-\delta t})\Delta\log(q) \right)$$
$$\approx \frac{\psi}{1-(1-\tilde{x}_0)e^{-\delta t}} \left(\frac{\delta}{\lambda-\delta}(e^{-\delta t} - e^{-\lambda t})\hat{q}_0 + (1-e^{-\delta t})\Delta\log(q) \right)$$

Notice that Proposition 2 implies that

$$\Delta\log(q) = \frac{1}{\chi}\left(\theta-1\right)\Delta\log A$$

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

$$\hat{q}_0 = \Delta \log(q_0) - \Delta \log(q) = \frac{1}{\chi}\frac{\lambda - \delta}{\delta}(\theta - 1)\,\Delta \log A$$

Thus,

$$\begin{aligned}
\log\left(\frac{e_t^{young}}{e_{0^-}^{young}}\right) - \log\left(\frac{e_t^{old}}{e_{0^-}^{old}}\right) &\approx \frac{\psi}{\chi}\frac{1}{1-(1-\tilde{x}_0)e^{-\delta t}}\left((e^{-\delta t} - e^{-\lambda t}) + (1 - e^{-\delta t})\right)(\theta - 1)\log A \\
&\approx \frac{\psi}{\chi}\frac{1 - e^{-\lambda t}}{1-(1-\tilde{x}_0)e^{-\delta t}}(\theta - 1)\Delta \log A,
\end{aligned}$$

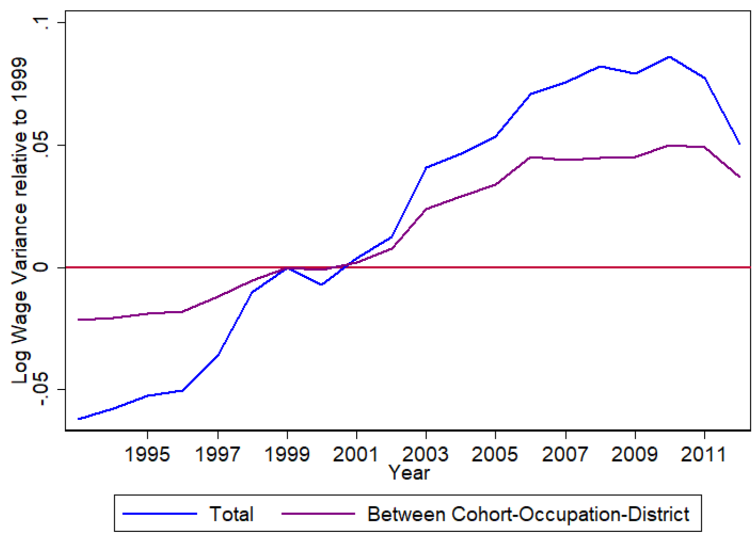
which is equivalent to (25).

Appendix B Empirical Application

B.1 Sample statistics

We begin with illustrating the increase in inequality, measured by the standard deviation of log wages, in our sample. Figure 9 compares the overall change in inequality together with the between district-cohort-occupation component, measured using the residual log-wage dispersion from a mincer regression including dummies for the district-cohort-occupation estimated on the sample in each year. Between 1997-2012, overall inequality in our sample increased by about 8.5 log points. As the figure illustrates, the between district-cohort-occupation component explains about half of the increase in inequality during this period. Separately, the characteristics do not account for the inequality rise (results available on request). Notice that the explanatory power of this component is similar to that of establishment dummies as showing in (Card, Heining, and Kline, 2013) as the main driver of the inequality increase in Germany during this period.[26]

Figure 9: Aggregate Trends in Log Wage Variance



Notes: Estimation of the aggregate standard deviation of log wages on the full LIAB sample and the residual dispersion in log wages from a mincer regression including district-occupation-generation dummies. Estimates are changes in dispersion relative to 1999.

Table A1 presents summary statistics underlying the FDZ microdata used in our empirical analysis, illustrating the evolution of the number of employees, ages and mean log wages of the baseline cohorts used in estimating (27)-(28).

[26] As there are nearly 50 times as many establishments as district-occupation-generation pairs in our sample, this is not mechanical.

Table A1: Summary Statistics: German Microdata

	1997	2014
Number of observations		
Born before 1960 (“Old”)	183,706	96,045
Born after 1960 (“Young”)	278,122	538,590
Mean log wage		
Born before 1960 (“Old”)	4.56	4.42
Born after 1960 (“Young”)	4.17	4.54
Mean age		
Born before 1960 (“Old”)	46.77	60.53
Born after 1960 (“Young”)	29.30	39.56

Notes. Sample of male workers in LIAB data, living in West Germany, employed full-time with a positive wage in 120 occupations. Cohorts as defined in the table.

B.2 Impact of New Technologies on Cognitive-intense Occupations

We next illustrate that the results in Table 1 and Figures 7-8 are robust to varying sample or generation definitions, as well as controls. We also present results for the aggregate wagebill both by the cognitive-intensity of the occupation and generation in Table A2.

As is clear from Table A2, wagebills in Germany experienced a significant increase in relatively more cognitive-intensive occupations (Panel A). This increase was driven by a large increase in the wagebills of the top third of cognitive-intensive occupations, as shown in Panel B of the table. The overall increase is evident for both young and old generations, though the relative increase is stronger for the young generation. For both generations, the aggregate increase is driven by an increase in the wagebills of the top third of cognitive-intensive occupations. For the old generation, this is offset by a relative decline in wagebills for the bottom third of the cognitive-intensive occupations.

Appendix Table A3 presents estimates from estimating (26) using (a) alternative controls, (b) alternative definitions of generations and (c) alternative sample periods. Panel A of the table illustrates that the results are robust to dropping the baseline controls of import and export exposure and the growth in the fraction of migrants in the occupation. The results are also robust to restricting the sample to native-born German males only. Panel B presents results where the "young" generation is defined alternatively as those born after 1965 or 1955, and also where the "young" generation are defined year by year as those aged below 40 in a year. In all cases, the results are similar to the baseline. As expected, when the definition of the young generation is further restricted relative to the baseline of being born in 1960, the coefficient on "Young" is stronger. Panel C illustrates that the results are not sensitive to varying the time period under analysis, though the overall effect is insignificant when considering only the period 1999-2014.

Appendix Tables A4 and A5 illustrate the robustness of the results from estimating (27)-(28). As is evident from Table A4, the results are similar when we vary the controls in the estimation. The three panels of Table A4 present estimates for the period before the shock (1996-1999, Panel A), the period during which DSL was rolled out across German regions (1999-2007, Panel B) and the entire post-shock period of the sample (1999-2014, Panel C). Each panel includes the results of our baseline specification, as well as alternative specifications where we include either no additional district-level, generation or occupation controls, and a specification where we drop the pre-trend control. In the absence of the pretrend control, the young generation has a significant pretrend in the pre-shock period, which results in a pretrend in the overall effect as well. In the absence of all controls, there are stronger pretrends in both generations, as well as in the between generation effect. This is exactly what we would expect as the initial determinants of the locations of the MDFs (our instrument) are likely correlated with long-run trends in labor market outcomes over regions. The addition of the controls clearly helps mitigate these correlations, and the pretrend control completely eliminates them. As we require that the instruments are exogenous conditional on the controls, this finding is reassuring. We note that, as panel Panel B and C highlight, the post-shock effects are positive for the young and negative and small for the old generation in all cases. In fact, controlling for the pretrend, as we do in the baseline estimates, results in the relative effect of the shock on the young generation being even stronger than in the other specifications.

Table A5 presents results relating to varying the definition of the young generation, as well as robustness to restricting the sample to only native-born males. We consider several definitions of the young generation – those born after 1955, 1965 or 1970, and generations

Table A2: Cognitive Intensity and Employment Growth in Germany

Dependent variable: Log-change in occupation payroll, 1997-2014	All (1)	Old (2)	Young (3)
Panel A: Linear specification			
Cognitive intensity	0.013***	0.009***	0.016***
	(0.003)	(0.003)	(0.003)
Panel B: Nonlinear specification			
Percentile of cognitive intensity			
Low: below percentile 30	-0.120	-1.28***	0.185
	(0.384)	(0.283)	(0.340)
Medium: percentiles 30-60	-0.053	0.060	-0.046
	(0.189)	(0.192)	(0.201)
High: above percentile 60	0.786***	0.599***	0.973**
	(0.207)	(0.221)	(0.212)

Notes. Sample of 120 occupations in 1997 and 2014. Young cohort defined as all workers born after 1960 and Old cohort as all workers born before 1960. All regressions include a set of baseline controls: growth in occupational exposure to exports from 1997-2014, growth in occupational exposure to imports from 1997-2014, and growth in the fraction of immigrants in the occupation in 1997-2014. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

defined by age in each year (ages less than 35, 40 or 45). In all cases we include the baseline controls discussed in Section 5. Once again, the results are similar across specifications – the coefficient on the cognitive intensity of the occupation for young workers is always positive and strongly significant, while that for the old generation is insignificant and close to zero. The differential effect between the young and old workers (the column labeled "Between") is always positive, though it is not significant in the cases where the young generation is defined very broadly to include workers who might more reasonably be thought of as being part of the older generation (for instance, those between 40 and 45 when DSL is rolled out across Germany).

Table A3: Robustness: Cognitive Intensity and Employment Growth in Germany

Dependent variable: Log-change in occupation employment			
	All	Young	Old
Panel A: Alternative control set			
No Controls	0.008***	0.011***	0.002
	(0.003)	(0.003)	(0.003)
Native-born Males Only	0.007**	0.011***	-0.001
	(0.003)	(0.004)	(0.003)
Panel B: Alternative cohort definition			
Young: Born after 1965	0.007**	0.013***	0.001
	(0.003)	(0.004)	(0.003)
Young: Born after 1955	0.007**	0.009**	-0.002
	(0.0037)	(0.004)	(0.004)
Young: Aged below 40	0.007**	0.011***	0.001
	(0.004)	(0.004)	(0.003)
Panel C: Alternative Sample Periods			
1993-2014	0.009**	0.014***	0.000
	(0.004)	(0.004)	(0.003)
1999-2014	0.004	0.007**	-0.003
	(0.003)	(0.004)	(0.004)

Notes. Sample of 120 occupations, sample periods as defined in the table. Table reports the estimated coefficient on the occupation's cognitive intensity in equation (26). Each row defines a separate robustness exercise. All regressions except "No Controls" include a set of baseline controls: growth in occupational exposure to exports during the sample period, growth in occupational exposure to imports during the sample period, and growth in the fraction of immigrants in the occupation during the sample period. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Effect of DSL Access Across Regions: Alternative Control Sets

Control Set	All	Young	Old	Between
Panel A: 1996-1999				
Baseline	-0.002	0.011	-0.019	0.029
	(0.026)	(0.030)	(0.031)	(0.049)
No Controls	-0.241***	-0.288***	-0.139**	-0.149**
	(0.068)	(0.082)	(0.058)	(0.064)
No Pretrend Control	-0.109*	-0.141*	-0.074	-0.068
	(0.065)	(0.077)	(0.084)	(0.061)
Panel B: 1999-2007				
Baseline	0.077*	0.223***	-0.138	0.361**
	(0.043)	(0.092)	(0.116)	(0.177)
No Controls	0.196*	0.272*	-0.159	0.431***
	(0.114)	(0.153)	(0.134)	(0.159)
No Pretrend Control	0.015	0.137	-0.200	0.337
	(0.061)	(0.085)	(0.127)	(0.149)
Panel C: 1999-2014				
Baseline	0.240***	0.482***	-0.065	0.546**
	(0.085)	(0.154)	(0.193)	(0.287)
No Controls	0.223	0.335	-0.224	0.559**
	(0.212)	(0.266)	(0.246)	(0.244)
No Pretrend Control	0.177**	0.292***	-0.026	0.319
	(0.087)	(0.114)	(0.189)	(0.222)

Notes. Sample of 120 occupations, sample periods as defined in the table. Column 1 reports the estimated coefficient on interaction between the occupation's cognitive intensity and district DSL access in equation (27). Column 2-3 reports the estimated coefficients on interaction between the occupation's cognitive intensity, a cohort fixed effect and district DSL access in equation (28), and their difference (column 4). Cohorts are the baseline cohorts with young workers those born after 1960. Each row defines a separate robustness exercise. All regressions except "No Controls" include a set of baseline district-level controls as well as occupation-year and cohort-year fixed effects.
Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Effect of DSL Access Across Regions: Robustness to Sample Selection

	Sample	All	Young	Old	Between
1999-2007					
	Baseline	0.077*	0.223***	-0.138	0.361**
		(0.043)	(0.092)	(0.116)	(0.177)
	Native-born Males Only	0.054	0.145*	0.037	0.108
		(0.045)	(0.092)	(0.105)	(0.171)
	Young: born after 1970		0.449***	-0.070	0.518*
			(0.120)	(0.191)	(0.289)
	Young: born after 1965		0.298***	-0.168	0.465**
			(0.092)	(0.118)	(0.183)
	Young: born after 1955		0.203**	-0.155	0.358***
			(0.087)	(0.115)	(0.167)
	Young: Aged < 35 in each year		0.118**	0.095	0.022
			(0.066)	(0.095)	(0.130)
	Young: Aged < 40 in each year		0.195**	0.030	0.165
			(0.086)	(0.104)	(0.166)
	Young: Aged < 45 in each year		0.206***	0.091	0.115
			(0.090)	(0.111)	(0.174)
1999-2014					
	Baseline	0.240***	0.482***	-0.065	0.546**
		(0.085)	(0.154)	(0.193)	(0.287)
	Native-born Males Only	0.223***	0.446***	0.074	0.372**
		(0.078)	(0.143)	(0.144)	(0.208)
	Young: born after 1970		0.714***	-0.048	0.789**
			(0.182)	(0.242)	(0.371)
	Young: born after 1965		0.612***	-0.171	0.783***
			(0.157)	(0.201)	(0.303)
	Young: born after 1955		0.573***	-0.298	0.871**
			(0.196)	(0.233)	(0.355)
	Young: Aged < 35 in each year		0.612***	0.059	0.553***
			(0.139)	(0.139)	(0.203)
	Young: Aged < 40 in each year		0.529***	0.076	0.453**
			(0.164)	(0.163)	(0.237)
	Young: Aged < 45 in each year		0.445***	0.159	0.286
			(0.170)	(0.198)	(0.266)

Notes. Sample of 120 occupations, sample periods as defined in the table. Column 1 reports the estimated coefficient on interaction between the occupation's cognitive intensity and district DSL access in equation (27). Column 2-3 reports the estimated coefficients on interaction between the occupation's cognitive intensity, a cohort fixed effect and district DSL access in equation (28), and their difference (column 4). Each row defines a separate sample selection exercise, alternatively restricting the baseline sample to only Germans ("Native-born") or varying the definition of young workers. All regressions include a set of baseline district-level controls as well as occupation-year and cohort-year fixed effects and a pre-trend control.
Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B.3 Calibration

This appendix discusses in detail the calibration of the model. As discussed in Section 6.1, we calibrate $\theta = 3$ and $\rho = 0.02$. We also select the distribution of innate ability to normalize the initial skill distribution to be uniform: $s_0(i) \equiv 1$. In addition, we abstract from differences across skills in non-cognitive productivity: $\alpha(i) \equiv 1$.

We now describe the procedure to calibrate all remaining parameters of the model using the estimates in Section 5. For all parameters, we assume that the shock starts with the roll-out of broadband internet in 2003. We then select parameters to match the estimates for the period of 2008 to 2014 in which we find statistically significant response in the relative payroll and relative employment of cognitive-intensive occupations.

Cohort size: δ and $\tilde{x}_0$. We first set $\tilde{x}_0$ to match the 60% share of young workers in the national population in 1997. We then select δ to match the incline of 25 p.p. in the share of young workers in population between 1997 and 2014. Specifically, we select x and δ such that

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

$$x = -\frac{1}{\delta} \log 0.4$$

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

Speed of Adjustment: λ. Proposition 2 implies that it is possible to write the impulse response function of relative output as

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

where $\alpha_0 > 0$, $\alpha_1 < 0$, and $\lambda > 0$.

We select the parameter λ to match the growth in the estimates response of relative payroll of more cognitive intensive occupations:

$$\hat{\lambda} = arg \min_{\lambda} \sum_{t=2008}^{2014} \left[\left(\hat{\beta}_t^y - \hat{\beta}_{2007}^y \right) - \alpha_1 e^{-\lambda(t-2007)} \right]^2 \tag{B.1}$$

where $\hat{\beta}_t^y$ is the estimated coefficient of (27) reported in Panel B of Figure 7.

The minimization problem in (B.1) yields $\hat{\lambda} = 0.135$. Figure 11 shows the fit of the calibrated model

Long-run skill supply elasticity: ψ. To calibrate ψ, we first construct the parameter

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

Our baseline calibration implies that $\hat{\alpha} = 3.484$.

Proposition 1 implies that

$$\kappa\eta = \psi\hat{\alpha} - \theta \tag{B.2}$$

Using expression (25), we have that

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

From Proposition 2,

$$(\theta - 1)\, \Delta log(A) = \Delta \log(y_t) \left(\frac{1 + \kappa\eta}{\theta + \kappa\eta} + \frac{\psi}{\chi} \frac{\theta - 1}{\theta + \kappa\eta} (1 - e^{-\lambda t}) \right)^{-1} \tag{B.3}$$

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

Combining these two expressions, we get that

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

Using the expression for $\kappa\eta$ in (B.2),

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

We then define the function:

$$F^{\psi}(\psi, t) \equiv \frac{1 - e^{-\hat{\lambda} t}}{1 - (1 - \tilde{x}_0) e^{-\hat{\delta} t}} \left((\rho + \hat{\delta})\, \frac{1 + \psi\hat{\alpha} - \theta}{\psi} + 1 - \frac{\theta - 1}{\psi\hat{\alpha}} e^{-\hat{\lambda} t} \right)^{-1}.$$

We select the parameter ψ to match the ratio of the between-generation employment response and the payroll response:

$$\hat{\psi} = arg \min_{\psi} \sum_{t=2008}^{2014} \left[\frac{\hat{\beta}_t^{young} - \hat{\beta}_t^{old}}{\hat{\beta}_t^{y}} - F^{\psi}(\psi, t) \right]^2 \tag{B.4}$$

where $\hat{\beta}_t^{y}$ is the estimated coefficient of (27) reported in Panel B of Figure 7, and $\hat{\beta}_t^{young} - \hat{\beta}_t^{old}$ is the between-generation employment response obtained from the estimation of (28) reported in Panel B of Figure 8.

The minimization problem in (B.4) yields $\hat{\psi} = 0.345$. Figure 11 shows the fit of the calibrated model.

Short-run skill supply elasticity: η. The combination of (24) and (B.3) implies that

$$\frac{\Delta \log e_t^{old}}{\Delta \log y_t} \approx \frac{\eta}{e_{H,0^{-1}}} \frac{1 - \frac{\psi}{\chi}(1 - e^{-\lambda t})}{1 + \kappa\eta + \frac{\psi}{\chi}(\theta - 1)(1 - e^{-\lambda t})}.$$

Using the expression for $\kappa\eta$ in (B.2),

$$\frac{\Delta \log e_t^{old}}{\Delta \log y_t} \approx \frac{\eta}{e_{H,0^-}} \frac{1 - \frac{(\theta-1)(1-e^{-\lambda t})}{\alpha(\rho+\delta)+1}}{1 + \psi\alpha - \theta + \frac{(\theta-1)(1-e^{-\lambda t})}{\alpha(\rho+\delta)+1}}.$$

We then define

$$F^{\eta}(\eta,t) \equiv \frac{\eta}{e_{H,0^-}} \frac{1-\frac{\theta-1}{\hat{\alpha}\left(\rho+\hat{\delta}\right)+1}(1-e^{-\hat{\lambda}t})}{1+\hat{\psi}\hat{\alpha}-\theta+\frac{\theta-1}{\hat{\alpha}\left(\rho+\hat{\delta}\right)+1}(1-e^{-\hat{\lambda}t})}.$$

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

We select the parameter η to match the ratio of the employment response of old workers and the payroll response:

$$\hat{\eta} = arg\min_{\eta} \sum_{t=2008}^{2014} \left[\frac{\hat{\beta}_t^{old}}{\hat{\beta}_t^{y}} - F^{\eta}(\eta,t)\right]^2 \tag{B.5}$$

where $\hat{\beta}_t^{y}$ is the estimated coefficient of (27) reported in Panel B of Figure 7, and $\hat{\beta}_t^{old}$ is the employment response for old workers obtained from the estimation of (28) reported in Panel A of Figure 8.

The negative point estimates reported in Panel A of Figure 8 imply that the minimization problem in (B.5) yields $\hat{\eta} < 0$. Since the employment response of old generations is small and nonsignificant, we assume that they are identical to zero, which yields $\hat{\eta} = 0$. Hence, we calibrate $\eta = 0$ and evaluate the model predictions under alternative specifications of this parameter.

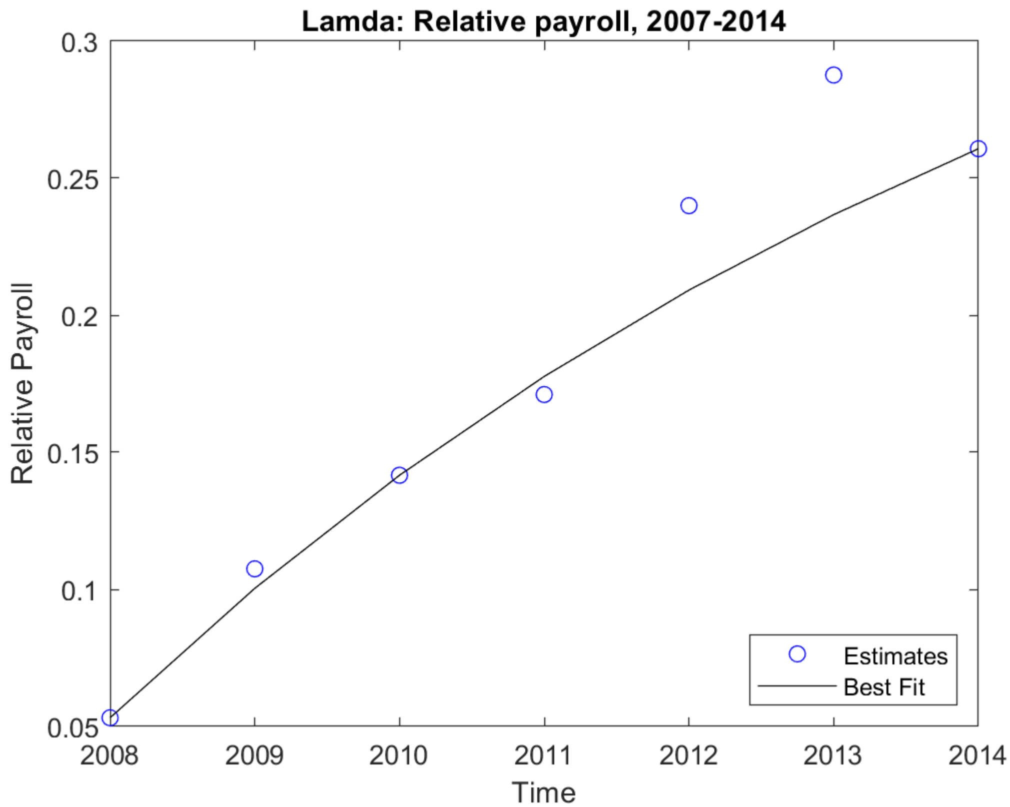


Figure 10: Calibration of λ

Notes: Blue dots represent the point estimates of β_t reported in Panel B of Figure 7. Black solid curve represents the bet fit line with $\lambda = 0.135$ obtained from the solution of (B.1).

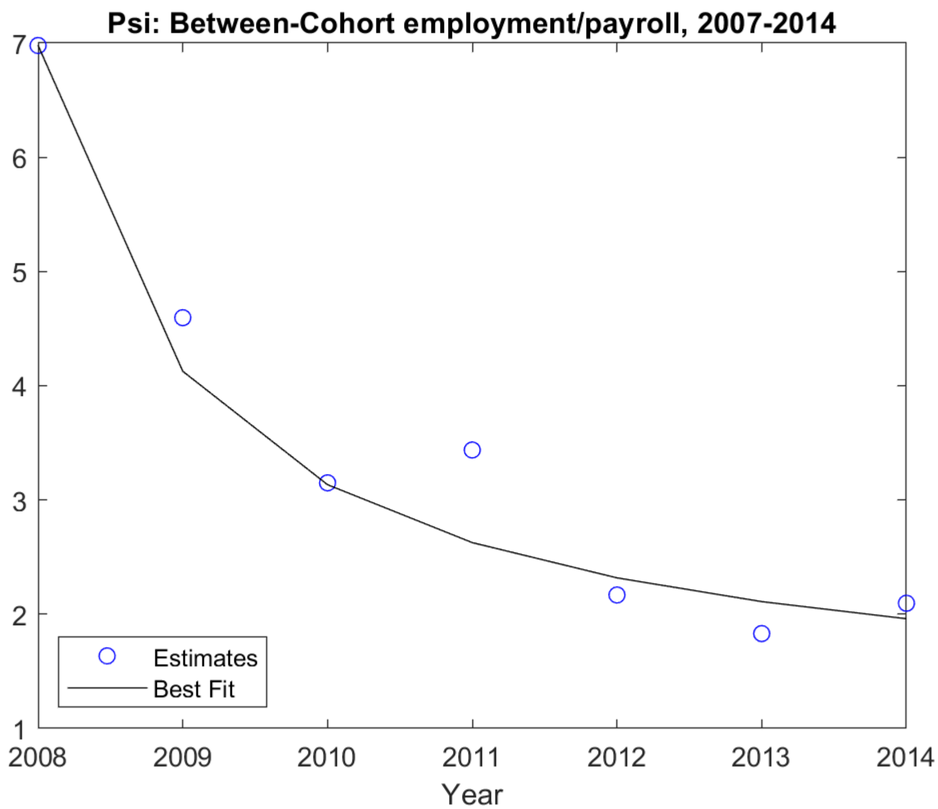


Figure 11: Calibration of ψ

Notes: Blue dots represent the point estimates of $\frac{\hat{\beta}_t^{young} - \hat{\beta}_t^{old}}{\hat{\beta}_t^{y}}$ using the estimates reported in Panel B of Figures 7 and 8. Black solid curve corresponds to $F^{\psi}(\hat{\psi}, t)$ with $\hat{\psi} = 0.354$ obtained from the solution of (B.4).

Appendix C Additional Theoretical Results

C.1 Extensions

TBA

C.2 Alternative microfoundations of the relative demand in (2)

TBA

C.3 Microfoundation of the intermediate good technology

Production Technology. The output of firm k at time t aggregates per-worker output $y_{kt}(i)$,

$$Y_{kt} = \int_0^1 y_{kt}(i) s_{kt}(i) di$$

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

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

$$y_{kt}(i) = a_C(i)^{\beta_k} a_{NC}(i)^{1-\beta_k},$$

where β_k denotes the production intensity of firm k on cognitive tasks.

In our model, *skill-technology specificity* arises whenever firms are heterogeneous in terms of task intensity and workers are heterogeneous in terms of skill bundle. To see this, suppose that firm C's technology uses cognitive tasks more intensely than firm NC's technology, $\beta_C > \beta_{NC}$, and that a worker of type i has a higher cognitive-noncognitive skill ratio than a worker of type j, $a_C(i)/a_{NC}(i) > a_C(j)/a_{NC}(j)$. In this case, i has a higher relative output with the cognitive-intensive technology than j, $y_{Ct}(i)/y_{NCt}(i) > y_{Ct}(j)/y_{NCt}(j)$, and, therefore, type i is more complementary to the cognitive-intensive technology than type j. As discussed below, worker-technology complementarity implies that, in equilibrium, workers with different skills sort to use different technologies and, therefore, are subject to different types of technological shocks.

In the rest of the paper, we assume the existence of such worker-technology complementarity and impose that the production process of firm $k = C$ is more intensive in cognitive tasks than that of firm $k = NC$. We also assume that types differ in terms of their skill bundle and, without loss of generality, impose that high-i types are relatively better in performing cognitive-intensive tasks.

Assumption 1 (Worker-technology complementarity)

1. *Firm C's technology uses cognitive tasks more intensely than firm NC's technology:* $\beta_C > \beta_{NC}$.

2. *High-i types have higher cognitive-noncognitive skill ratio:* $\sigma(i) \equiv \left(\frac{a_C(i)}{a_{NC}(i)}\right)^{\beta_C - \beta_{NC}}$ *is differentiable and strictly increasing in i.*