TECHNOLOGICAL TRANSITIONS WITH SKILL HETEROGENEITY ACROSS
GENERATIONS

Rodrigo Adão
Martin Beraja
Nitya Pandalai-Nayar

ABSTRACT

Why are some technological transitions particularly unequal and slow to play out? We develop a theory to study transitions after technological innovations driven by worker reallocation within a generation and changes in the skill distribution across generations. The economy's transitional dynamics have a representation as a q-theory of skill investment. We exploit this in two ways. First, to show that technology-skill specificity and the cost of skill investment determine how unequal and slow transitions are by affecting the two adjustment margins in the theory. Second, to connect these determinants to measurable, short-horizon changes in labor market outcomes within and between generations. We then empirically analyze the adjustment to recent cognitive-biased innovations in developed economies. Strong responses of cognitive-intensive employment for young but not old generations suggest that cognitive-skill specificity is high and that the supply of cognitive skills is more elastic for younger generations. This evidence indicates that cognitive-biased transitions slowly unfold over many generations. As such, naively extrapolating from observed changes at short horizons leads to overly pessimistic views about their welfare and distributional implications.

Rodrigo Adão
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
rodrigo.adao@chicagobooth.edu

Nitya Pandalai-Nayar
Department of Economics
University of Texas at Austin
3.142 BRB, 2225 Speedway
Austin, TX 78713
and NBER
npnayar@utexas.edu

Martin Beraja
Department of Economics
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
martinberaja@gmail.com

An online appendix is available at https://www.nber.org/data-appendix/w26625/

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] If an economy's adjustment margins vary as horizons lengthen, then focusing on short or long horizons alone risks missing the overall impact of technological innovations on labor markets as well as their average and distributional welfare consequences. Such concerns are particularly important when the adjustment is slow and takes many generations. How then do economies adjust to technological innovations over different horizons? Why are some technological transitions particularly unequal and slow to play out?

In this paper, we develop a theory to study technological transitions driven by both worker reallocation within a generation and changes in the distribution of skills across generations. We show that the transitional dynamics of this economy can be represented as a q-theory of skill investment. We then use this representation in two ways. First, to characterize how the nature of the technological innovation and associated skills determine the importance of the two adjustment margins in the theory and, as a result, how slow and unequal technological transitions are. Second, to connect these determinants to changes in labor market outcomes within and between generations that can be measured at short horizons. We conclude by empirically studying the adjustment of developed economies to recent cognitive-biased technological innovations. Our results indicate that, due to the high specificity of cognitive skills, such innovations triggered a particularly unequal and slow transition unfolding over many generations. In such cases, by missing most of the adjustment, naive extrapolations from observed changes at short horizons overestimate the distributional consequences and underestimate the average welfare gains of cognitive-biased innovations.

The theory has four 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 the relative technology-specific wage at a point in time, there is a threshold determining which skill types self-select into each of the two technologies. *Technology-skill specificity* —i.e., how different skill types are in terms of their relative technology-specific productivity—then determines how sensitive the assignment threshold is to changes in relative technology-specific wages. Third, the output of the two technologies is combined to produce a final consumption good, as in Katz and Murphy (1992), Ngai and Pissarides (2007), and Buera, Kaboski, and Shin (2011). Fourth, given future relative technology-specific wages, workers make a costly investment upon entering the labor market that determines their skill type for their lifetime, similar to Chari and Hopenhayn (1991), Caselli (1999) and Galor and Moav (2002). The *cost*

---

[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 reviews of the literature documenting the impact of new technologies on employment and wages of workers associated with different skills, occupations, industries, and firms.

of skill investment for entering workers then determines how different the skill distribution is across generations following changes in future relative technology-specific wages.[2]

The equilibrium of this economy is a joint path for the skill distribution, the assignment of skill types to technologies, and the relative technology-specific wage and output. It entails a high-dimensional fixed-point problem: forward-looking entrants make skill investment decisions based on the expected future path for the relative technology-specific wage, which determines how the skill distribution evolves over time and, ultimately, the actual equilibrium path of relative technology-specific wage and all other outcomes.

Our first result reduces the dimensionality of this fixed-point problem. It establishes that the approximate equilibrium of this economy can be represented as that from a q-theory of skill investment.[3] The path for the skill distribution is only a function of two variables at each point in time: the present-discounted value of the log-relative technology-specific wage (q) and the threshold determining the assignment of skills to technologies (which plays the role of the pre-determined variable). We show that a simple system of linear differential equations characterizes the equilibrium dynamics of these two variables. Thus, we solve for the equilibrium dynamics by keeping track of these variables and not the skill distribution itself. Our approach is reminiscent of those in Perla and Tonetti (2014) and the special case with linear objectives in Lucas and Moll (2014) which characterize the dynamics of a distribution by tracking the evolution of a threshold.

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 relative labor demand and wages in the improved technology. On impact, marginal skill types within each generation reallocate into that technology. The increase in current and future relative wages leads younger entering generations to invest in those skills that are more complementary to the improved technology. Along the transition, as younger generations replace older ones, q falls and relative output increases because the economy's skill distribution tilts towards skills more complementary to the improved technology. To evaluate how slow the transition is, we define the discounted cumulative impulse response (DCIR). For old generations born before the innovation, the DCIR measures how different is the adjustment they expect to see during their lifetime compared to the overall (long-run) adjustment. We say that the adjustment is slower whenever they expect to miss more of the overall adjustment (i.e., the DCIR is smaller). Crucially, we also show that the DCIR of q is a central determinant of the average and distributional welfare consequences of new technologies.

---

[2]We incorporate investment on the continuum of skills by allowing ex-ante identical individuals to enter a skill lottery whose utility cost is proportional to the relative entropy between the chosen lottery and a reference distribution of innate ability. This formulation yields a tractable expression for the skill distribution of incoming generations that resembles a multinomial logit function over the continuum of skill types.

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

This result shows that the impact of new technologies on the economy may significantly change over time due to the endogenous evolution of the skill distribution across generations. It provides a micro-foundation for the idea that supply elasticities tend to be lower at shorter horizons compared to longer horizons, a form of Samuelson's LeChatelier principle. Our micro-foundation points to two types of risks associated with ignoring dynamics induced by changes in the skill distribution across generations. The first arises when extrapolating from observed responses in the economy that span much less than a generation. Such extrapolations will overestimate inequality changes and underestimate average welfare gains. The second arises when extrapolating from past technological transitions to different contexts: a type of threat to external validity. This leads to biased predictions about the economy's dynamic adjustment whenever the nature of technology and skills, or the underlying flexibility of skill investment, significantly differs across episodes.

Our third result presents comparative static exercises that speak to why some technological transitions are particularly unequal and slow to play out. As such, these exercises help interpret differences between past or future transitions where the nature of technological innovations and associated skills differ. First, we show that an economy where technology-skill specificity is higher has a slower, more back-loaded adjustment path to the new long-run equilibrium. The q-theory analogy again delivers the intuition for this result. When technology-skill specificity is higher, there is less worker reallocation across technologies in the short-run and, therefore, the increase in lifetime inequality q is larger. This strengthens the incentives of young entering generations to invest in those skills that are more complementary to the improved technology. As a result, the adjustment is slower because larger changes in the skill distribution take place along the transition. Second, we show that a lower cost of skill investment for young generations makes the adjustment slower as well, both directly and by amplifying the effects of technology-skill specificity. In both cases, the more relevant margin of adjustment is not the reallocation of workers within a generation but the changes in the skill distribution across generations.

Our fourth result connects the degree of technology-skill specificity and the cost of skill investment to observable changes in labor market outcomes within and between generations. In particular, we focus on short-run implications that can be credibly measured in most datasets. Our measurement insight is that, in the short-run, economies with higher technology-skill specificity are associated with weaker within-generation changes in the relative employment of older workers across occupations (or sectors), but stronger between-generation differences in the relative employment of younger and older workers. In contrast, a lower cost of skill investment for entering generations is also associated with larger between-generation differences in relative employment, but has no effect on the responses for older workers. Such generation-specific changes are common in empirical analysis of how economies adjust to different types of shocks.[4] By connecting them to structural parameters, our theory shows how

---

[4]For example, Kim and Topel (1995), Card and Lemieux (2001), Autor and Dorn (2009), Autor, Dorn, and Hanson (2013), Guvenen et al. (2017), McCaig and Pavcnik (2018), Greenland, Lopresti, and McHenry (2019), Porzio and Santangelo (2019).

these measurable moments in the short-run are informative about the economy's transitional dynamics and, consequently, how unequal and slow the adjustment will be.

In the second part of the paper, we explore our theory's observable predictions to provide three pieces of evidence indicating that technology-skill specificity and changes in the skill distribution across generations are relevant to understand how developed economies adjusted to recent cognitive-biased technological innovations.[5] First, we analyze employment trends in nine broad occupation groups in eighteen developed countries. We document that, in all countries, employment growth in the three most cognitive-intensive occupations was stronger for younger workers than for older workers. Second, we use microdata to provide a more detailed investigation of these responses in Germany. Controlling for a number of confounding factors, we show that employment and payroll grew more in occupations that require more time spent performing cognitive-intensive tasks. We find that the effect of cognitive intensity on these variables is strong for younger generations, but weak for older generations. We also explore the unique large-scale German training program to document higher growth in the number of trainees in more cognitive intensive occupations, suggesting that incoming cohorts tilt investment towards cognitive skills. Finally, following Falck, Gold, and Heblich (2014), we use pre-determined conditions of the German telephone network to obtain quasi-experimental variation across regions in the adoption timing of broadband internet in the early 2000s. By comparing late to early adopting regions, we estimate impulse response functions that show an increase in the relative employment and payroll of more cognitive-intensive occupations starting in 2005. The estimates are again different for older and young generations. The impact on relative employment is small and nonsignificant for older generations, but it is positive and statistically significant for younger generations.

In sum, this evidence suggests that cognitive-skill specificity is high and that the supply of cognitive skills is elastic at longer horizons. Parameterizing our model to match the empirical impulse responses for Germany, we find that these two features make cognitive-biased transitions particularly unequal and slow. As a result, we quantify that, compared to naively extrapolating from observed changes on impact, the true average welfare (lifetime welfare inequality) increase across generations is about 50 percent higher (lower) following a relatively large cognitive-biased innovation. Had technology-skill specificity been lower, such innovation would have led to a faster and less unequal transition instead, featuring smaller between-generation differences in occupation composition changes. Compared to most recent years, we document that these between-generation differences were smaller in the United States and Germany in the 1960s and 1970s, which is consistent with the nature of technological innovations and skills being different in the two periods.

---

[5]Our approach follows an extensive literature documenting that the recent arrival of new technologies in the workplace, like the computer and the internet, augmented the productivity of jobs intensive in cognitive tasks while substituted jobs intensive in routine tasks —e.g. Autor, Levy, and Murnane (2003), Spitz-Oener (2006), Autor and Dorn (2013), Akerman, Gaarder, and Mogstad (2015), Acemoglu and Restrepo (2017), and, for a review, Acemoglu and Autor (2011).

Related literature.   Our paper is related to several strands of the literature. A long literature has analyzed structural transformation in the form of long-run worker reallocation across sectors – e.g., Ngai and Pissarides (2007), Buera, Kaboski, and Shin (2011), Buera and Kaboski (2012), 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 between-generation differences in employment reallocation across sectors following long-run changes in sectoral productivity growth (Kim and Topel (1995) and Hobijn, Schoellman, and Vindas (2019)) or schooling (Porzio and Santangelo (2019)). We make two contributions to this literature. First, we provide a tractable theory to analyze how skill heterogeneity within and across generations shapes the transitional dynamics induced by technological innovations. This allows us to characterize how fast the transition is, a focus we share with Gabaix, Lasry, Lions, and Moll (2016). We use this characterization to point out which features of the economy lead to slow adjustment dynamics and large biases from welfare calculations that ignore them. Second, we estimate impulse response functions to a technological innovation in Germany and show how they discipline our theoretical mechanisms.

The only source of dynamics in our theory is the endogenous change in the distribution of skills across generations. This mechanism is consistent with recent evidence documenting the impact of labor demand shocks on young individuals' decisions of educational attainment (Atkin (2016) and Charles, Hurst, and Notowidigdo (2018)) and field of study (Abramitzky, Lavy, and Segev (2019), Ghose (2019) and, for a review, Altonji, Arcidiacono, and Maurel (2016)). We add to this literature by documenting that cognitive-biased innovations differentially affect employment and training in cognitive-intensive occupations across generations in Germany. Several papers have proposed alternative sources of dynamics to study technological transitions, 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), endogenous creation of new tasks for labor in production (Acemoglu and Restrepo, 2018), mobility costs of heterogeneous workers (Dvorkin and Monge-Naranjo, 2019), and rising wealth inequality via permanent changes in the returns to wealth 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 affects the economy's adjustment to skill-biased technological innovations.

An extensive 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). 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 technology-skill specificity and skill investment costs for younger generations generate slower adjustments following skill-biased innovations. We then link the two adjustment margins in our theory to observable responses of labor market outcomes within and between generations. Our empirical application indicates that separately allowing for these two forces is important in the context of the recent experiences of developed countries, in general, and Germany, in particular.

Our paper is also related to the literature that has estimated the distributional consequences of shocks to the demand and supply of skills – for a review, see Acemoglu and Autor (2011). Our empirical analysis follows the literature studying the impact of new technologies on the demand of skills across occupations with different task intensity – e.g., Autor, Levy, and Murnane (2003), Autor and Dorn (2013) and Acemoglu and Restrepo (2017). As Akerman, Gaarder, and Mogstad (2015), we estimate the labor market consequences of broadband internet adoption. While they focus on the impact on educational composition of employment in Norwegian firms, we estimate its effect on the occupation composition of employment across German local labor markets. Similar to Autor and Dorn (2009), we find that the impact of new technologies differs for younger and older workers. Relative to this literature, our results indicate that such reduced-form evidence estimated at short horizons is informative about structural parameters governing the dynamic adjustment to new technologies, but they also caution against directly extrapolating from it when technological transitions are slow.[6]

Finally, we note that a full account of inequality trends in developed countries would require analyzing not only the impact of shocks to skill demand but also how the economy adjusts to exogenous shocks to the supply of skills, as in Katz and Murphy (1992) and Goldin and Katz (2009). For instance, Card and Lemieux (2001) show that generation-specific skill supply shocks are important determinants of inequality trends. As in their paper, we incorporate skill differences across generations. However, while it would be possible to study the consequences of skill supply shocks in our theory, we focus on the consequences of shocks to skill demand when skill supply endogenously responds across generations.

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 how skill-biased technological innovations affect welfare and labor market outcomes over different horizons.

---

[6]Heckman, Lochner, and Taber (1998) point out that such extrapolations may be biased because prices and skills adjust in general equilibrium. This motivated a quantitative literature studying the labor market consequences of sector- or occupation-specific demand shocks – e.g., Lee and Wolpin (2006), Kambourov and Manovskii (2009), Dix-Carneiro (2014).

Sector 4 shows how technology-skill specificity and the skill investment cost determine how slow and unequal the transition is. Section 5 links the main ingredients of our theory to responses in observable outcomes for different generations of workers in the short-run. In Section 6, we empirically analyze the adjustment of developed economies to recent cognitive-biased innovations. Section 7 presents our quantitative analysis. Section 8 concludes.

2 A Model of Skilled-biased Technological Transitions

2.1 Environment

We consider a closed economy in continuous time. There is a single final good 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, $i \in [0, 1]$. The skill type determines the worker's productivity with each production technology.

Final good. Production of the final good is a CES aggregator of the two inputs:

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

where $\theta > 0$ is the demand elasticity of substitution between the low-tech and the high-tech intermediate inputs, and A_t is a shifter of the relative productivity of the high-tech input (as in Katz and Murphy (1992)).

Conditional on input prices, the cost minimization problem of firms producing the final good implies that the relative spending on the high-tech input is

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

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

We consider a competitive environment where zero profit determines the final good price,

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

Assignment of skills to technologies. 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 their 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, the labor income of a worker with skill type i is

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

The technology-skill assignment in equation (7) plays a central role in determining the economy's adjustment to technological shocks. Equation (7) illustrates that such an assignment depends on the endogenous price ω_t defining the relative value of one unit of effective labor employed in high-tech production, as well as the exogenous function $\sigma(i)$ defining the differential productivity of type i in high-tech production. Without loss of generality, we assume that $\sigma(i)$ is increasing: we order types such that higher i types have higher relative productivity in high-tech production. Recent papers have considered a similar structure of endogenous sorting of workers to different technologies – e.g., Acemoglu and Autor (2011), Costinot and Vogel (2010), Adão (2016).

In our theory, ω_t is a natural measure of inequality as it is the endogenous relative wage rate of skill types employed in different technologies conditional on their productivity. In what follows, we will refer to ω_t as the relative technology-specific wage or, sometimes, simply as the relative wage. However, it is important to notice that movements in ω_t are not perfectly aligned with movements in the relative labor income of high-tech employees. As pointed out by Heckman and Honore (1990), the endogenous assignment problem in (7) implies that high-tech relative labor income may change due to changes in the "selection" of skill types employed in high-tech production – that is, changes in the average $\sigma(i)$ and $\alpha(i)$ of types employed with the H technology. Adão (2016) shows that, depending on the shape of $\alpha(i)$, these selection forces may amplify or offset the impact of ω_t on the average income of high-tech employees.

Skill investment. We now endogeneize the distribution of skills by allowing workers to direct their skill investment decisions to target particular skill types. After we define the competitive equilibrium, we discuss the economic interpretation of our main assumptions as well as why they yield a mathematically tractable theory of technological transitions.

We consider an overlapping generations setting in which the birth and death of workers follows a Poisson process with rate δ.[7] 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

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

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 future path for the wage distribution $\{w_s(i)\}_{s>t}$, 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 exogenous distribution of innate ability, $\bar{s}_t(i)$. A worker's type is then fixed during their lifetime.[8]

Formally, we assume that the cost of the lottery is proportional to the Kullback-Leibler divergence between the lottery $\tilde{s}_t(i)$ and the baseline distribution $\bar{s}_t(i)$, so that workers of the cohort born at time t solve the following skill investment problem:

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

The positive parameter ψ governs the cost of targeting particular skill types. In the limit when $\psi \to 0$, the cost of targeting a particular skill type 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, $V_t(i)$.

Equilibrium. The assumption that only new generations choose skill lotteries implies that the evolution of the skill distribution $s_t(i)$ is given by the following 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 that relative demand and supply of the high-tech good are equal:

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

where y_t is given by (2) and x_t is the ratio of high- to low-tech production given by (4)–(5).

---

[7]In Section 4.4, we consider an extension where population grows because the birth rate is higher than the death rate.
[8]In Section 4.4, we extend the model to allow workers to re-optimize later in life following following a technological innovation.

Definition 1 (Competitive Equilibrium) *Given an initial skill distribution $s_0(i)$ and exogenous paths for $\{A_t, \bar{s}_t(i)\}_{t \geq 0}$, a competitive equilibrium is a path of the technology-skill assignment $\{G_t(i) : i \in [0,1] \rightarrow \{H, L\}\}_{t \geq 0}$, the skill distribution $\{s_t(i)\}_{t \geq 0}$, the skill lottery $\{\tilde{s}_t(i)\}_{t \geq 0}$, the relative value of output $\{y_t\}_{t \geq 0}$, the relative wage and final price index $\{\omega_t, P_t\}_{t \geq 0}$, such that*

1. *Given ω_t, the technology-skill assignment is given by the self-selection decisions in (7).*

2. *Given $\{\omega_t\}_{t \geq 0}$, the skill lottery is given by the skill investment decisions of new cohorts in (9).*

3. *Given $s_0(i)$ and $\{\tilde{s}_t(i)\}_{t \geq 0}$, the skill distribution follows the Kolmogorov-Forward equation (10).*

4. *The final price index is given by (3).*

5. *For all $t \geq 0$, the technology-skill assignment, the skill distribution, the relative value of output, and the relative wage satisfy the market clearing condition in (11).*

Discussion. A number of comments on the assumptions and their economic interpretation are in order. There are admittedly four strong assumptions that we make for simplicity and tractability. The first is that there is a continuum of skill types that determine the productivity in each technology. The second is that $\bar{s}_t(i)$ is exogenous. The third is that only new incoming generations can invest in skills in response to changes in relative wages. Old generations may only respond to such wage changes by moving across technologies but not by changing their skill-type. The fourth is that skill investments have an uncertain outcome represented by the skill lottery $\tilde{s}_t(i)$ whose cost takes the particular functional form in (9).

First, we assume the existence of a continuum of skills for two main reasons. As discussed below, this assumption implies that changes in the technology-skill assignment are smooth along the transition – that is, any relative wage change triggers the re-allocation of a positive mass of skill types. Second, as discussed in Acemoglu and Autor (2011), Roy-like skill heterogeneity yields responses in worker allocations and wages following technological innovations that do not arise in the canonical model with skills specified in terms of observable worker attributes. Such heterogeneity then allows us to study, for example, how technology-skill specificity affects the economy's adjustment.

Regarding a skill-type's economic interpretation, in Appendix E.1 we provide a microfoundation of (4)–(5) where the production of each good combines individual-level output given by a Cobb-Douglas function of each worker's "cognitive" and "non-cognitive" task input. Production of the H good is more intensive in cognitive tasks, with $\sigma(i)$ denoting the H-to-L output ratio of i's differential ability to perform cognitive tasks. Thus, to observe a type i in this setting, it is necessary to know each worker's ability to perform cognitive and non-cognitive tasks. We assume that workers and firms observe such abilities, but treat them as unobservable to researchers.

The second and third assumptions imply that the flow of new workers to a particular point in the skill distribution is independent of the current skill distribution (as can be seen from equation (10)). This simplifies the law of motion of the skill-distribution and allows us to characterize its dynamics in general equilibrium. The independence results from the

fact that skill investment decisions are independent from a worker's current skill-type (since only young workers invest and are born without a type) and because $\bar{s}_t(i)$ is exogenous. In Appendix B, we relax both assumptions. First, we endogenize $\bar{s}_t(i)$ by considering an extension where workers can "learn from others." Specifically, we assume that $\bar{s}_t(i)$ depends on the current skill distribution in a way that makes it is easier for workers to target skills that are already abundant. Second, we allow old generations to re-optimize their skill investments as well. We show that what is important for our main results in the coming sections is that the cost of skill investment is lower for younger generations when compared to that for older generations. It is not essential that this cost is infinite for older generations, as in our baseline specification.

Our preferred economic interpretation of these baseline assumptions is that changes in relative wages may induce older workers in the labor market to switch towards sectors or occupations that require similar skills and may thus entail minimal re-training. To fundamentally change career paths by acquiring completely different skills, however, they may face a high cost. For younger workers though, such skill investments are less costly due to lower opportunity cost or higher ability to learn new skills.[9] For tractability, we collapse these investments that in reality occur either through formal schooling or on-the-job into a one-time decision upon entering the market.

Regarding the fourth assumption, we make it purely for reasons of tractability. Different from theories of uni-dimensional human capital investment, workers in our theory can direct their investments to target specific skill types. Yet, mathematically, this *directed skill investment* problem is in principle substantially more complex. As we will see below, assuming that ex-ante identical workers choose an uninsurable lottery by incurring the entropy-based cost delivers a tractable problem with a non-degenerate skill distribution as a solution. What is important for our results though is not that workers are ex-ante identical,[10] but (again) the independence of the skill investment from a worker's current skill-type.[11] As for the entropy cost function, it has a long tradition in macroeconomics.[12] As discussed later, this function yields the continuous time analog of the optimal skill investment in an environment in which worker's ability to acquire a discrete number of skills follows a Type 1 extreme-value distribu-

[9]In line with this interpretation, Lee and Wolpin (2006) show that older workers exhibit much lower mobility across occupations and sectors than younger workers.

[10]Note that it is straight forward to extend our model to introduce ex-ante heterogeneity in observed worker attributes that only affect $\bar{s}(i)$ and ψ in the optimal skill investment problem in (9). The overall skill lottery in this extension is simply the average of lotteries across worker-groups with different attributes. Such an extension does not affect our main theoretical insights, but affects the mapping of the ingredients of the model to the data.

[11]Such assumptions would be violated if the type of an individual entering the skill lottery affects the relative cost or benefit of particular lotteries. This is the case, for example, if there is inter-generation transmission of skills, or skill acquisition has monetary costs in an environment with credit frictions.

[12]For example, it has been used to compute the distance between distributions in frameworks with rational inattention (Sims (2003)) and model uncertainty (Hansen and Sargent (2008)). Note, however, that the fact the cost is proportional to the relative entropy is not as important for our main results, since in later sections we log-linearize the path for all variables around the stationary equilibrium. Instead, what matters is the curvature of the distance metric around the stationary equilibrium, similar to investment problems with a convex cost of adjustment.

tion.[13] In our theory though, having a continuum of skill types is useful when combined with continuous time because it implies that the dynamic adjustment of all outcomes is smooth along the equilibrium path. This helps us sharply characterize the economy's transitional dynamics in general equilibrium.

Our preferred interpretation for the uncertainty of the skill type realization is that individuals with different unobservables may have heterogeneous returns to education and on-the-job training. This is in fact consistent with the evidence in Carneiro, Heckman, and Vytlacil (2011). Our model treats this heterogeneity in unobservables through the uncertainty of the type realization.

2.2 Static and dynamic equilibrium conditions

We now proceed to derive equilibrium conditions in two steps. First, we consider static conditions that, given the skill distribution $s_t(i)$ at time t, determine the technology-skill assignment, the relative wage ω_t, and the relative output value y_t. Second, we consider dynamic conditions that, given the path of the relative wage $\{\omega_t\}_{t\geq 0}$, determine the optimal skill lottery chosen by entering generations $\{\tilde{s}_t(i)\}_{t\geq 0}$ and thus the evolution of the skill distribution $\{s_t(i)\}_{t\geq 0}$.

Static equilibrium conditions. The endogenous sorting decision in (7) determines the assignment of skill types to technologies. It 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. Since $\sigma(i)$ is increasing, the assignment is described by a threshold l_t characterizing the type that is indifferent between working with any of the two technologies. The following lemma formalizes this discussion.

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*

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

Lemma 1 links the relative wage ω_t to the allocation of skill types across technologies. Condition (12) is central to understand the impact of technological shocks on the allocation of workers across technologies. The slope of $\sigma(l_t)$ determines the strength of the *comparative advantage* in high-tech production of skill types slightly below l_t compared to that of skill type l_t. Thus, as shown by Acemoglu and Autor (2011) and Costinot and Vogel (2010), it essentially determines how much the relative wage must increase to induce the reallocation of skill types below l_t from the L to the H technology. Accordingly, the inverse elasticity of $\sigma(i)$ controls the

---

[13]With a discrete number of skill types, our specification yields skill choices that are isomorphic to those implied by a discrete-choice problem a la McFadden et al. (1973). It can thus be seen as a generalization of this framework when there is a continuous of available choices.

subset of skill types that reallocate across technologies in response to changes in the relative wage. Formally, (12) implies that

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

where $l_t(\omega_t)$ is the implicit function defined by (12). Since the economy's skill distribution does not adjust instantaneously, the inverse elasticity of $\sigma(i)$ plays the role of the short-run elasticity of skill supply across technologies. Then, whenever the elasticity of $\sigma(i)$ (i.e., $1/\eta$) is higher, we will say that the degree of *technology-skill specificity* is higher because the induced short-run worker reallocation following relative wage changes is smaller.

The technology-skill assignment in Lemma 1 together with equations (4)–(5) imply that the relative supply of high-tech production is

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

The threshold l_t is then determined by the market clearing condition in (11). Whenever l_t is high, equation (12) implies that ω_t is low and, therefore, the relative demand for good H is high. In this case, however, the relative high-tech supply is low as only a small share of types are employed in high-tech production. Whenever l_t is low, the opposite is true. In equilibrium, there is a single threshold that equalizes relative demand and supply given the skill distribution $s_t(i)$.

Lemma 2 (Equilibrium Threshold) *Given $s_t(i)$ and A_t, there is a unique equilibrium threshold l_t that guarantees goods market clearing,*

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

Proof. See Appendix A.1. ∎

Dynamic equilibrium conditions. We now turn to the entrant's forward-looking problem of choosing their skill lottery $\tilde{s}_t(i)$ given the path of the relative wage $\{\omega_s\}_{s>t}$. The solution of the maximization problem in (9) yields the following optimal lottery.

Lemma 3 (Optimal Lottery) *Define $\log(Q_t(i)) \equiv \int_t^\infty e^{-(\rho+\delta)(s-t)} \max\{\log(\omega_s\sigma(i)), 0\}ds$. The optimal lottery is*

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

Proof. See Appendix A.2. ∎

The optimal lottery in (15) is a multinomial logit function over 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 $Q_t(i)$. 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 $\log(W(i)) = \frac{\log(\alpha(i)\max\{\omega\sigma(i),1\})}{\rho+\delta}$ is the present discounted log-wage of skill type i.

In this case, the skill distribution is a constant-elasticity function of relative wages across types, where the elasticity is ψ.[14] 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, ψ governs the long-run skill supply across technologies, which we formally define as

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

In the rest of the paper, we refer to $1/\psi$ as the *cost of skill investment*, which is inversely related to the long-run skill supply across technologies.

2.3 Skill distribution dynamics: A q-theory of skill investment

We now combine the static and dynamic equilibrium conditions to solve for the equilibrium path of the skill distribution as well as all other equilibrium variables, given an arbitrary initial skill distribution $s_0(i)$ and a constant path for $\{A_t, \bar{s}_t(i)\}_{t\geq0}$.

In principle, this involves solving a complex infinite-dimensional fixed-point problem. To see this, consider a conjectured path for the relative wage $\{\omega_t\}_{t\geq0}$. This path determines the skill investment decisions of new generations in (15) and, as such, the path for the skill distribution $\{s_t(i)\}_{t\geq0}$ from (10) given $s_0(i)$. The relative wage path also determines the assignment threshold path ($\{l_t\}_{t\geq0}$) from the indifference condition (12). Taken together, the skill distribution and the assignment threshold determine the relative supply of the high-tech input ($\{x_t\}_{t\geq0}$). In an equilibrium, the relative supply of the high-tech input needs to be equal to its relative demand at the the conjectured path for the relative wage – i.e., they need to be consistent with market clearing.

Our first result approximates the solution of this fixed-point problem by considering a log-linear expansion around the stationary equilibrium. It establishes that the approximate equilibrium of this economy can be represented as that of a q-theory of skill investment, where q refers to the present discounted value of the log-relative wage or, as we call it from

---

[14]Notice that the long-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 the stationary skill distribution.

now on, the relative lifetime wage:

$$log(q_t) \equiv \int_t^\infty e^{-(\rho+\delta)(s-t)} log(\omega_s)ds.$$

Specifically, we show that one does not need to keep track of the whole skill distribution to solve for the equilibrium path of the relative lifetime wage, q_t, and the assignment threshold, l_t. The equilibrium dynamics of these two variables are fully characterized by a simple system of linear differential equations. Letting " ˆ " denote variables in log-deviations from the stationary equilibrium, the following theorem presents the system of differential equations that, given $\hat{l}_0$, determines the equilibrium path of $\{\hat{q}_t, \hat{l}_t\}_t$ when $\{A_t, \bar{s}_t(i)\}_{t\geq0}$ are constant over time.[15] It also characterizes the skill distribution $\hat{s}_t(i)$, the optimal lottery $\hat{\tilde{s}}_t(i)$ and relative value of output $\hat{y}_t$, given the equilibrium path of $\{\hat{q}_t, \hat{l}_t\}_t$ and the initial distribution $s_0(i)$.

Theorem 1 (A q-theory of skill investment) *Suppose that $\{A_t, \bar{s}_t(i)\}_{t\geq0}$ are constant over time.*

1. *Given initial condition $\hat{l}_0$ and terminal condition $\lim_{t\to\infty} \hat{l}_t = 0$, the equilibrium dynamics of $\{\hat{q}_t, \hat{l}_t\}$ are described by the system of differential equations*

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

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

where $\kappa > 0$ is a constant.

2. *The equilibrium $\{\hat{q}_t, \hat{l}_t\}_{t\geq0}$ is saddle-path stable and given by*

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

where

$$\lambda = -\frac{\rho}{2} + \sqrt{\left(\frac{\rho}{2}\right)^2 + \delta\left((\rho+\delta) + \frac{\psi}{\theta+\kappa\eta}\right)} \quad and \quad \zeta = -\frac{1}{\eta}\frac{1}{\rho+\delta+\lambda}. \tag{20}$$

3. *The equilibrium dynamics of the skill distribution $\hat{s}_t(i)$, the optimal lotteries $\hat{\tilde{s}}_t(i)$, and the value of relative high-tech output $\hat{y}_t$ are given by*

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

$$\hat{s}_t(i) = \hat{s}_0(i)e^{-\delta t} + \int_0^t e^{\delta(\tau-t)}\hat{\tilde{s}}_\tau(i)d\tau, \tag{22}$$

$$\hat{y}_t = (\theta - 1)\frac{1}{\eta}\hat{l}_t, \tag{23}$$

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

Proof. See Appendix A.3. ∎

The first part of the theorem presents a system that is a rather standard one in macroeconomics. The assignment threshold, $\hat{l}_t$, is a state variable whose law of motion needs to be solved backward. The present discounted value of relative technology-specific wages, $\hat{q}_t$, is a control variable whose law of motion needs to be solved forward. The system is in fact mathematically isomorphic to the q-theory of capital investment (Tobin, 1969, Hayashi, 1982). In our model, $\hat{q}_t$ is the present discounted value of the relative wage in high-tech production, representing the shadow price of the human capital "asset" associated with having one additional unit of the high-tech good. Whenever this price is higher, the incentives to invest in high-i skills are stronger. As in the seminal q-theory, parameters governing the costs of adjustment in the economy (i.e., δ and ψ) affect the sensitivity of changes in the assignment threshold $\frac{\partial \hat{l}_t}{\partial t}$ to $\hat{q}_t$. However, our model features both imperfect substitution of human capital across technologies and heterogeneous skills. Thus, the impact of q_t on the evolution of l_t also depends on the degree of technology-skill specificity (as measured by η) and the substitutability of inputs (as measured by θ).

The other two parts of the theorem characterize the equilibrium dynamics of various outcomes. The second part shows that (locally) the equilibrium exists and is unique—a consequence of saddle-path stability. Given an initial condition $\hat{l}_0$, both $\hat{l}_t$ and $\hat{q}_t$ converge at a constant rate of λ to the stationary equilibrium. The expressions in (19) show that, whenever $\hat{l}_0 < 0$, l_t increases and q_t decreases along the equilibrium path. The last part of the theorem links the equilibrium path of the skill distribution and the relative high-tech output to the joint dynamics of $\{\hat{q}_t, \hat{l}_t\}$. The change in the optimal skill lottery in (21) along the transition depends centrally on the evolution of the relative return of skills employed in high-tech production as measured by $\hat{q}_t$. The parameter ψ controls the sensitivity of the optimal skill investment to such changes. The overall skill distribution in (22) 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 value of relative high-tech output is driven by changes in relative wages, $\hat{y}_t = (\theta - 1)\hat{\omega}_t$. The sensitivity of such wage changes to changes in the threshold depends on the degree of technology-skill specificity controlled by η.

Theorem 1 reduces the dimensionality of the equilibrium's fixed-point problem. It characterizes the equilibrium dynamics of $\{\hat{q}_t, \hat{l}_t\}_{t \geq 0}$ without tracking the full skill distribution $s_t(i)$. This is possible for three reasons. First, the dynamics of $s_t(i)$ only depend on $\log(Q_t(i))$ via the optimal skill lottery. Yet, $\log(q_t)$ suffices to determine the value of most skill-types in this skill investment decision—as opposed to the full path of relative wages ω_t in $\log(Q_t(i))$—because most workers never switch technologies along an equilibrium path whenever relative wages are close to their stationary level. Second, the market clearing condition (14) determining l_t only contains integrals of $s_t(i)$. Because of the continuum of skill types, the effect of the marginal types that switch technologies are of second order when evaluating changes in these

integrals. Taken together, the two observations imply that changes in $\hat{l}_t$ over time are a function of $\hat{q}_t$ and $\hat{l}_t$, as described in (17). Finally, since the indifference condition (12) yields a mapping between ω_t and l_t, the dynamics of $\hat{q}_t$ can be written as a function of the future path of l_t, as described by (18).

3   The Adjustment to Skill-biased Technological Innovations

We now analyze the dynamic adjustment of our economy to a permanent, unanticipated 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. We use the results from the previous section to characterize in closed-form the dynamic responses of q_t, l_t and y_t, as well as the evolution of the skill distribution $s_t(i)$. The dynamic responses indicate that the impact of new technologies on the economy may significantly change over time due to the endogenous evolution of the skill distribution across generations. How slow this adjustment is then crucially determines the average and distributional welfare consequences of new technologies.

Our results show that the economy's adjustment is shaped by a form of Samuelson's LeChatelier principle: the elasticity of relative output supply increases over time due to changes in the skill distribution across generations. The dynamics of this relative supply elasticity points to the risk of using *naive* reduced-form representations of the economy to extrapolate from observed changes for short horizons or different historical contexts. Our theory formally establishes a micro-foundation for the idea that the relative supply elasticity changes over time. In the rest of the paper, we use this micro-foundation to study the determinants of skill-biased technological innovations, and to recover them from observed short-run responses of labor market outcomes following technological innovations.

3.1   Dynamic responses of equilibrium outcomes

We assume that immediately prior to the shock at time $t = 0^-$ the economy is in a stationary equilibrium. We let $\Delta \log(A) > 0$ be the relative productivity shock, and denote log-changes in outcomes as $\Delta \log(q_t) \equiv \log(q_t/q_{0^-})$, $\Delta \log(y_t) \equiv \log(y_t/y_{0^-})$, and $\Delta \log(l_t) \equiv \log(l_t/l_{0^-})$.

Proposition 1 (Dynamic responses) *Given a skill-biased technological innovation $\Delta \log(A)$, the dynamic responses $\Delta \log(l_t)$, $\Delta \log(q_t)$ and $\Delta \log(y_t)$ are approximated by:*

$$
\begin{bmatrix} \Delta \log(l_t) \\ \Delta \log(q_t) \\ \Delta \log(y_t) \end{bmatrix} = \underbrace{\begin{bmatrix} -\eta \\ \frac{1}{\rho+\lambda} \\ 1+\kappa\eta \end{bmatrix} \frac{\theta-1}{\theta+\kappa\eta} \Delta \log(A)}_{\text{Short-run}} + \underbrace{\frac{\psi}{\chi} \begin{bmatrix} \eta \\ \frac{-1}{\rho+\delta+\lambda} \\ \theta-1 \end{bmatrix} (1-e^{-\lambda t}) \frac{\theta-1}{\theta+\kappa\eta} \Delta \log(A)}_{\text{Short- to long-run transition}} \quad (24)
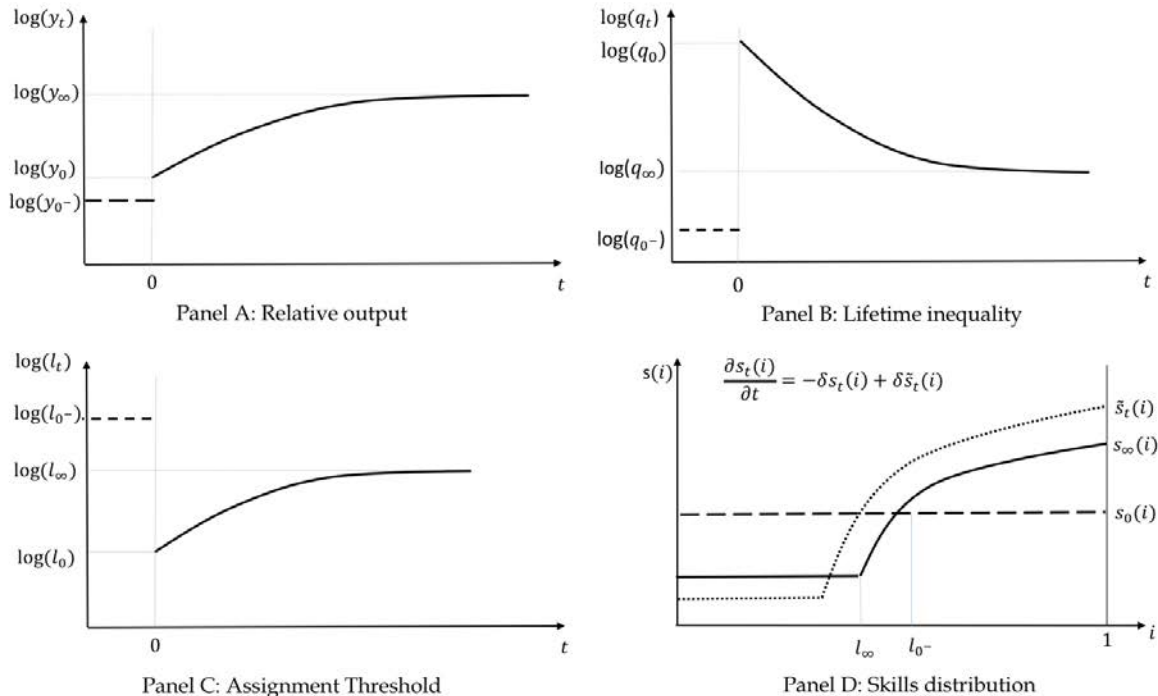$$

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

Proof. See Appendix A.4. ■

Figure 1 illustrates these dynamic responses together with the dynamics of the skill distribution and lottery. We do so for the case where the two technologies are substitutes in production ($\theta > 1$) and $\alpha(i) = 1$. The first term in (24) is the immediate impact of the shock represented by the responses at $t = 0$ in Figure 1. In the short-run, there are increases in both relative output ($\Delta \log(y_0) > 0$) and relative lifetime wages ($\Delta \log(q_0) > 0$). The higher relative wage in the H technology induces the reallocation of skill types in the existing worker generations from the L to the H technology, as can be seen from the decline in the assignment threshold l_t, which in turn adds to the increase in relative output.

The second term in (24) describes the change in all variables along the transition. It shows that these variables converge at the constant rate λ. The increase in the relative lifetime wage in high-tech production causes entering worker generations to twist their skill lotteries $\tilde{s}_\tau(i)$ towards high-i types whose skills are more complementary to high-tech production, as illustrated in the bottom left panel of Figure 1. This triggers changes in the economy's skill distribution $s_t(i)$ as older generations are replaced with younger generations at rate δ. Along the transition, the growing mass of high-i types employed with the H technology implies a continuing process of relative output increase and relative wage decline. The rising relative high-tech output implies that the consumption price index in (3) falls along the transition. By reducing the relative high-tech wage, the arrival of more high-i types in younger generations triggers the displacement of marginal i types from high-tech production over time, as illustrated by the increasing l_t in Figure 1.

Figure 1: The economy's adjustment to a skill-biased technological shock ($\theta > 1$)

Panel A: Relative output

Panel B: Lifetime inequality

Panel C: Assignment Threshold

Panel D: Skills distribution

18

In the long-run, the economy features higher relative wage and output in high-tech production, and a larger mass of workers in the high-tech sector. The increase in high-tech employment is driven both by a skill distribution with higher mass in high-i types and a lower assignment threshold of skill types employed in technology H. It is important to notice that the only source of dynamics in our theory is the skill investment decision of cohorts born after the shock. Whenever incoming cohorts cannot invest in skills (i.e., $\psi = 0$), the transitional dynamics term in (24) disappears and the responses in the long-run are identical to those observed in the short-run.

We conclude this section by defining the discounted cumulative impulse response (DCIR). It conveniently summarizes the importance of transitional dynamics and thus relates to how slowly economies adjust to skill-biased innovations. Intuitively, it is the answer to the question: from the point of view of generations alive just before the innovation happens, how different is the adjustment they expect to see during their lifetime compared to the overall (long-run) adjustment? We will then say that the economy's adjustment is *faster* when existing generations expect to see more of the overall adjustment during their lifetime (i.e., the DCIR is smaller). Instead, the adjustment is *slower* when these generations expect to miss more of it (i.e., the DCIR is larger).

Definition 2 (Discounted Cumulative Impulse Response) *For any variable z_t and innovation $\Delta log(A)$, the discounted cumulative impulse response $DCIR(z)$ is:*

$$DCIR(z) = \left| \int_0^\infty \delta e^{-\delta t} \frac{\Delta \log(z_t)}{\Delta \log(A)} dt - \frac{\Delta \log(z_\infty)}{\Delta \log(A)} \right|.$$

Formally, the DCIR is the distance between the long-run response and the expected response of $\log(z_t)$ during the initial generations' lifetime, since all generations born before the shock have exponentially distributed death probabilities with rate δ. This is a natural and convenient measure of the importance of transitional dynamics in our context for a number of reasons. First, it encodes not only the rate of convergence λ but also other relevant features of the impulse responses like how front-loaded they are. For instance, one could have an adjustment where the short- and long-run changes are almost identical—implying a DCIR close to zero—but the rate of convergence λ from the short- to the long-run is very low. According to the DCIR, we would intuitively say that it is a fast adjustment since almost all of the overall adjustment is completed on impact, whereas looking at λ alone would give the impression that the adjustment is slow. Second, the DCIR does not mechanically scale with the replacement rate of generations. If δ is higher, this mechanically increases λ (making the adjustment faster) but it also decreases the expected lifetime of a generation (meaning they expect to miss more of the adjustment). Finally, in the next section, we show that this measure of how slowly economies adjust is the relevant one for analyzing the welfare consequences of skill-biased innovations.

3.2 Changes in average welfare and lifetime welfare inequality

We now use the characterization of the economy's transitional dynamics above to compute the welfare consequences of skill-biased technological innovations. Our welfare measure is the ex-ante expected utility of individuals born at each point in time, which is given by the solution of the utility maximization problem in (9). Given the log-utility assumption, the consumption-equivalent welfare gain is the change in the ex-ante utility multiplied by $(\rho + \delta)$. Equations (7) and (9) imply that the consumption-equivalent utility of cohort τ is

$$U_\tau \equiv (\rho + \delta) \int_0^1 \tilde{s}_\tau(i) \left[\log\left(\alpha(i)^{\frac{1}{\rho+\delta}} Q_\tau(i) \right) - \int_\tau^\infty e^{-(\rho+\delta)(t-\tau)} log(P_t) \, dt - \frac{1}{\psi} \log\left(\frac{\tilde{s}_\tau(i)}{\bar{s}(i)} \right) \right] di,$$

where $\tilde{s}_\tau(i)$ is the skill distribution of cohort τ, $Q_\tau(i)$ is the present-discounted value of $\max\{\log(\omega_t \sigma(i)), 0\}$ defined in Lemma 3, and P_t is the ideal price index defined in (3).

To obtain an average welfare measure, we take an utilitarian approach by considering a weighted average of the ex-ante utility of different cohorts where cohort τ's weight is $re^{-r\tau}$ —as in Calvo and Obstfeld (1988) and Itskhoki and Moll (2019).[16] The average welfare is

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

To obtain a measure of welfare inequality, notice that the relative wage is the only endogenous component of the relative earnings of skill types employed in different technologies. Thus, we define lifetime welfare inequality for cohort τ as the consumption-equivalent of the present discounted value of the relative wage, $(\rho + \delta) \log q_\tau$. We again aggregate different worker cohorts by defining the economy's average lifetime welfare inequality as

$$\bar{\Omega} \equiv (\rho + \delta) r \int_0^\infty e^{-r\tau} \log(q_\tau) d\tau.$$

The following proposition characterizes the induced changes in average welfare $\Delta \bar{U} \equiv \bar{U} - U_{0^-}$ and lifetime welfare inequality $\Delta \bar{\Omega} \equiv \bar{\Omega} - \log(q_{0^-})$.

Proposition 2 (Average welfare and lifetime welfare inequality) *The changes in average welfare $\Delta \bar{U}$ and lifetime inequality $\Delta \bar{\Omega}$ are approximately:*

$$\Delta \bar{U} = \frac{y_\infty}{1 + y_\infty} \Delta \log(A) - \left(\frac{y_\infty}{1 + y_\infty} - \frac{e_\infty}{1 + e_\infty} \right) \Delta \bar{\Omega}$$

$$\Delta \bar{\Omega} = (\rho + \delta) \left(\Delta \log(q_\infty) + \frac{\lambda r}{r + \lambda} \int_0^\infty \hat{q}_\tau d\tau \right)$$

where $e_\infty \equiv \left(\int_{l_\infty}^1 s(i) di \right) / \left(\int_0^{l_\infty} s(i) di \right)$ is the relative high-tech employment in the long-run.

---

[16]The parameter r captures the idea that the welfare of future generations may be discounted at each point in time. To see this, it is useful to consider two extreme cases. When $r \to \infty$, the social welfare function completely ignores the welfare of all future cohorts. In the other extreme, when $r \to 0$, the social welfare function only gives positive weight to generations born in the new stationary equilibrium.

Proof. See Appendix A.5. ∎

First, consider the change in lifetime welfare inequality. It trivially increases when the long-run change $\Delta \log(q_\infty)$ is larger, but also when the overall magnitude of the response along the transition is larger, as measured by the cumulative impulse response $\int_0^\infty \hat{q}_\tau d\tau$. Furthermore, the relative importance of the cumulative impulse response is increasing in λ since it governs how fast lifetime inequality decays, and as such, between-generation differences in lifetime inequality.

Second, the proposition shows that the innovation causes average welfare to increase, but this increase is partially offset by higher welfare inequality whenever the H technology has a higher average wage than the L technology (i.e., $y_\infty > e_\infty$).[17] This offsetting effect of inequality arises because there are two consequences of the increase in the relative high-tech wage (i.e., $\Delta \log q_t > 0$). It increases the average wage in the economy by increasing the wage of those employed in high-tech production – a share $\frac{e_\infty}{1+e_\infty}$ of workers. However, it also raises the economy's price index by increasing the consumption cost of the high-tech good – a share $\frac{y_\infty}{1+y_\infty}$ of overall output. The negative impact of the price index on the average real wage dominates whenever the high-tech output share is higher than its employment share.

Importantly, Proposition 2 shows that the average and distributional welfare consequences of new technologies depend crucially on how slow the adjustment is. To see this, note that when $r = \delta$ then $\Delta \bar{\Omega} = (\rho + \delta)(\Delta log(q_\infty) + DCIR(q)\Delta log(A))$. Whenever the increase in the relative supply of H goods happens mostly through changes in the distribution of skills as opposed to worker reallocation within a generation, then such increases will take many generations to materialize. This implies that the average wage increase will be back-loaded and come far in the future, whereas the inequality increase will be front-loaded (i.e., $DCIR(q)$ larger). In other words, generations born before the shock expect to see a small fraction of the long-run increase in relative output and decline in the price index during their lifetimes, but expect to experience larger changes in inequality compared to the long-run. As a result, the increase in average welfare (lifetime welfare inequality) will be smaller (larger) in such economies.

3.3 LeChatelier Principle and the risks of extrapolation

In this section, we connect the adjustment predicted by our theory to that implied by a reduced-form supply and demand framework. For each t, our model implies that the relative output and wage in high-tech production solves the following system of equations:

$$\Delta \log(x_t) = (\theta - 1)\Delta \log(A) - \theta \Delta \log(\omega_t), \tag{25}$$

$$\Delta \log(x_t) = \varphi_t \Delta \log(\omega_t). \tag{26}$$

The first expression is the "relative demand equation" in (2). As discussed above, it is the

---

[17]Such a case arises if absolute advantage is positively correlated with the comparative advantage to operate the H-technology – i.e., $\alpha(i)$ is increasing in i.

cornerstone of the canonical model in Katz and Murphy (1992) and its extensions reviewed by Acemoglu and Autor (2011). The demand equation relates changes in relative demand, $\Delta \log(x_t)$, to changes in relative productivity, $\Delta \log(A)$, and relative wages, $\Delta \log(\omega_t)$. The parameter θ is the elasticity of substitution between the output of skill types employed in different technologies.

The second expression is the "relative supply equation" linking changes in relative output supply, $\Delta \log(x_t)$, to changes in relative wages, $\Delta \log(\omega_t)$. The parameter φ_t is the elasticity of relative output supply, which is a function of the degree of skill-technology specificity and the cost of adjusting skill investment. Specifically, we show in Appendix A.6 that the impulse response functions in Proposition 1 yield

$$\varphi_t = \frac{\kappa \eta \chi + \theta \psi (1 - e^{-\lambda t})}{(\theta + \kappa \eta)(\delta + \rho) + \psi e^{-\lambda t}}, \tag{27}$$

where $\varphi_t > 0$ and $\frac{\partial \varphi_t}{\partial t} \geq 0$ for all t.

The positive and increasing φ_t over time implies a form of Samuelson's LeChatelier principle: the relative supply of high-tech output is more elastic over longer horizons due to changes in the skill distribution over time. This is the main difference between our theory and the canonical model of Katz and Murphy (1992) in which relative supply does not respond to changes relative wages ($\varphi_t = 0$).

The positive elasticity arises from two sources in our framework. First, even if skills are exogenous ($\psi = 0$), the relative supply elasticity is positive because a fraction of the heterogeneous workers in the economy decides to reallocate across technologies in response to changes in the relative wages, as in Acemoglu and Autor (2011). Second, the change in the skill investment decision of cohorts born after the shock introduces an additional adjustment margin for relative supply. This margin becomes stronger over time as younger cohorts replace older cohorts, driving φ_t upwards along the transition – in fact, $\frac{\partial \varphi_t}{\partial t} > 0$ if, and only if, $\psi > 0$.

By microfounding the dynamics of the relative supply elasticity, our results point to two types of risks associated with using reduced-form estimates of φ_t. The first arises when extrapolating from observed responses in the economy over any given horizon. Consider a researcher who knows θ and obtains φ_T and $\Delta \log A$ from the estimated impact of a technological shock on relative output and wages at horizon T. Suppose this researcher then uses her estimates to analyze the consequences of skill-biased innovations. It is clear that the time-varying nature of the reduced-form parameter φ_t implies that predictions will be biased for any period other than T. Specifically, the researcher's predictions will overestimate (underestimate) inequality changes and underestimate (overestimate) relative output changes for any period after (before) horizon T. The researcher will also obtain biased estimates of the welfare consequences of the shock as she will wrongly conclude that the change in lifetime welfare inequality is $(\rho + \delta)\Delta \log q_T = \Delta \log \omega_T$, which may be higher or lower than $\bar{\Omega}$ depending on

the estimation horizon T.[18]

The second type of risk arises when extrapolating from past technological transitions in different contexts: a type of threat to external validity. Consider a researcher that has estimates of the full path for the reduced-form elasticity φ_t obtained in a particular historical episode. Suppose this researcher uses such estimates to make predictions about the dynamic consequences of a new technological innovation in a different economy or historical context. Our theory shows that if either technology-skill specificity (η) or the cost of skill investment (ψ) are different, then the path of φ_t will be different as well. Thus, the researcher will obtain biased predictions about the economy's adjustment at all horizons whenever the nature of technology and skills or the underlying flexibility of skill investment are significantly different across episodes.

4   Determinants of Skill-biased Technological Transitions

In this section, we analyze how parameters governing technology-skill specificity and the cost of skill investment affect the economy's adjustment to a skill-biased technological innovation. The comparative static exercises speak to when is it that technological transitions are more unequal and slower, with the adjustment mainly driven by changes in the skill distribution across generations as opposed to fast reallocation of workers within a generation. As such, they help interpret differences between historical episodes or future transitions where the nature of technological innovations and associated skills differ. Furthermore, describing which type of economies adjust more slowly helps understanding the welfare consequences of technological innovations and when is it that researchers should be more cautious when extrapolating from reduced-form elasticities estimated using observations at short horizons.

4.1   Comparative statics with respect to technology-skill specificity

Consider first how the economy's dynamic responses change with the degree of technology-skill specificity, i.e., with how different skill types are in terms of their relative productivity in the high-tech sector. In our theory, technology-skill specificity is inversely related to the short-run skill supply elasticity η. Thus, this exercise speaks to differences in the dynamics across episodes in which skills of incumbent workers were more or less easily transferable for use in the new improved technology.

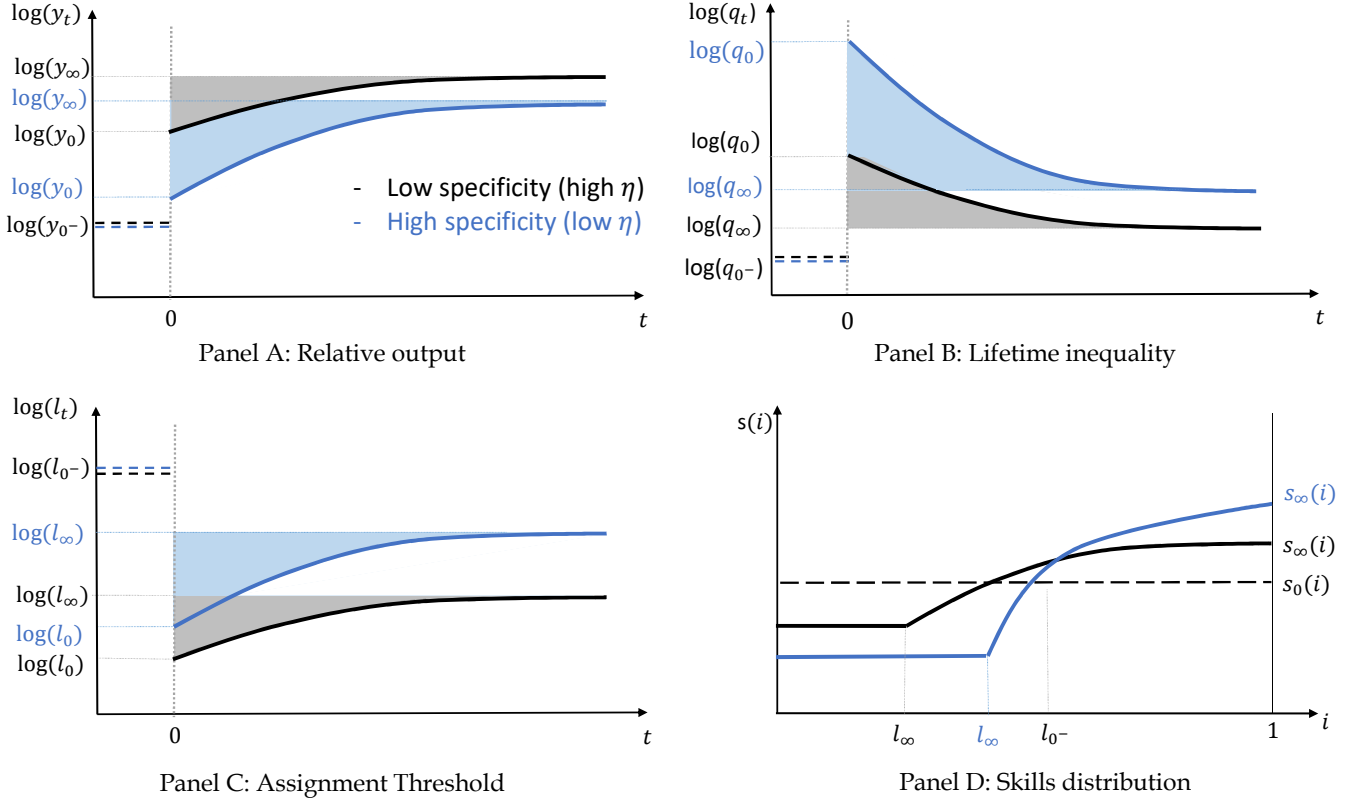Figure 2 shows the responses of two economies that differ in their degree of technology-

---

[18]To see this, consider the bias in the extreme case of $T = 0$, which yields the highest possible value of $\Delta \log \omega_0$:

$$\Delta \log \omega_0 - \Delta \bar{\Omega} = \left(1 + \frac{\rho + \delta}{r + \lambda} \right) \lambda^2 \int_0^\infty \hat{q}_t dt > 0.$$

In this case, the bias is proportional to the cumulative impulse response of inequality since an overall larger magnitude of the transitional dynamics amplifies the difference between the relative wage at $t = 0$ and the present discounted value of the change in lifetime inequality, $\Delta \bar{\Omega}$. In Appendix E.2, we consider researchers taking alternative approaches that ignore the dynamics of the skill distribution across generations.

skill specificity.[19] The black lines show the responses of an economy with a high value of η (i.e., low technology-skill specificity). The blue lines show the responses of an economy with a low value of η (i.e., high technology-skill specificity). In Appendix A.7, we support the graphical representation in Figure 2 with Proposition A.1 which establishes how η affects the short- and long-run responses, the cumulative impulse response, and the rate of convergence.

Figure 2: Comparative statics with respect to η

Panel A: Relative output

Panel B: Lifetime inequality

Panel C: Assignment Threshold

Panel D: Skills distribution

In the short-run, when technology-skill specificity is higher (lower η), a smaller mass of workers reallocate across technologies in response to the shock (i.e., $\frac{\partial |\Delta \log(l_0)|}{\partial \eta} > 0$). As a result, the short-run increase in relative wages and lifetime inequality q are larger (i.e., $\frac{\partial |\Delta \log(q_0)|}{\partial \eta} < 0$) and the increase in relative output is smaller (i.e., $\frac{\partial |\Delta \log(y_0)|}{\partial \eta} > 0$). The larger increase in q then implies that younger entering generations have stronger incentives to invest in those skills that are more complementary to the H technology. As a consequence, there are larger differences in skill heterogeneity across generations.[20] Then, the overall magnitude of the adjustment of y_t and q_t that happens along the transition is larger as well because larger changes in the skill distribution (the only slow moving state variable) take place as younger generations replace older generations. Formally, we measure this as the cumulative impulse response function being larger (e.g. $\frac{\partial \int_0^\infty \hat{q}_t dt}{\partial \eta} < 0$). Graphically, it corresponds to the blue

---

[19]The figure shows the case where $\theta > 1$ and the threshold's cumulative impulse response increases with η.

[20]This follows directly from the fact that $\tilde{s}_\tau(i)$ is proportional to $\hat{q}_\tau$ in (21) and $\hat{q}_\tau$ is larger for all τ when technology-skill specificity is higher.

shaded areas being larger than the black shaded areas.

Moreover, while the larger changes in the skill distribution could have implied a larger (smaller) overall long-run adjustment in relative output (lifetime inequality), it turns out that the smaller (larger) short-run response dominates. Thus, the long-run adjustment in relative output (lifetime inequality) is smaller (larger) in the economy with higher technology-skill specificity (i.e., $\frac{\partial |\Delta \log(y_\infty)|}{\partial \eta} > 0$, $\frac{\partial |\Delta \log(q_\infty)|}{\partial \eta} < 0$).

Finally, we can come back to the DCIR to summarize how technology-skill specificity affects the importance of transitional dynamics.

Theorem 2.1 (DCIR comparative statics with respect to η) *Following a skill-biased technological innovation $\Delta \log(A)$, lifetime inequality (q) and relative output (y) adjust slower in economies with a higher degree of technology-skill specificity (i.e., lower η). Formally,*

$$\frac{\partial DCIR(q)}{\partial \eta} < 0, \quad \frac{\partial DCIR(y)}{\partial \eta} < 0.$$

Proof. We have that $DCIR(q) = \frac{\lambda \delta}{\lambda + \delta} \left| \frac{\int_0^\infty \hat{q}_t dt}{\Delta \log(A)} \right|$. From Proposition A.1 in Appendix A.7, we know that when η is higher then the rate of convergence λ and the cumulative impulse response $\left| \int_0^\infty \hat{q}_t dt \right|$ are both smaller. The proof for y_t is analogous. ∎

The theorem shows that the DCIR is larger and transitional dynamics are more important in economies with a higher degree of technology-skill specificity. That is, the adjustment is slower and more back-loaded, with generations alive before the shock expecting to see less of the long-run changes during their lifetime. Intuitively, this is because the muted reallocation of workers at shorter horizons *causes* larger endogenous changes in the skill distribution along the transition due to the larger increases in lifetime inequality.[21]

It is also worth noting that the slower adjustment in economies with higher technology-skill specificity does not mechanically follow from the fact that reallocation is smaller in the short-run, neither from the fact that old generations are replaced slowly at rate δ. Instead, it follows from the skill distribution responding to stronger changes in relative wages. To make this point clear, Proposition 3 shows that technology-skill specificity has no effect on the DCIR of q and y when either the cost of skill investment for young generations is large ($\psi \to 0$) or the high- and low-tech inputs have a large elasticity of substitution and thus the relative wage responds little to technological innovations ($\theta \to \infty$).

Proposition 3 (Interaction of technology-skill specificity with ψ and θ)

1. *Small changes in the skill distribution ($\psi \to 0$)*

$$\left. \frac{\partial DCIR(y)}{\partial \eta} \right|_{\psi \to 0} = \left. \frac{\partial DCIR(q)}{\partial \eta} \right|_{\psi \to 0} = 0,$$

---

[21]More generally, after a shock, economies with a less mobile *stock* of a factor experience stronger changes in the *flow* of entrants because of larger changes in relative prices. For example, when old vintages of physical capital are less adaptable to use in a new sector, then the flow of firm entrants with newer capital vintages will be larger.

2. Small changes in inequality ($\theta \to \infty$)

$$\left.\frac{\partial DCIR(y)}{\partial \eta}\right|_{\theta \to \infty} = \left.\frac{\partial DCIR(q)}{\partial \eta}\right|_{\theta \to \infty} = 0$$

Proof. See Appendix A.8. ∎

To summarize, the results in this section highlight that how economies adjust to technological innovations depends crucially on the degree of technology-skill specificity. When technology-skill specificity is higher, technological transitions will be driven more by changes in the skill distribution across generations than the reallocation of workers within a generation. As a result, they will be more unequal and play out slower over many generations.
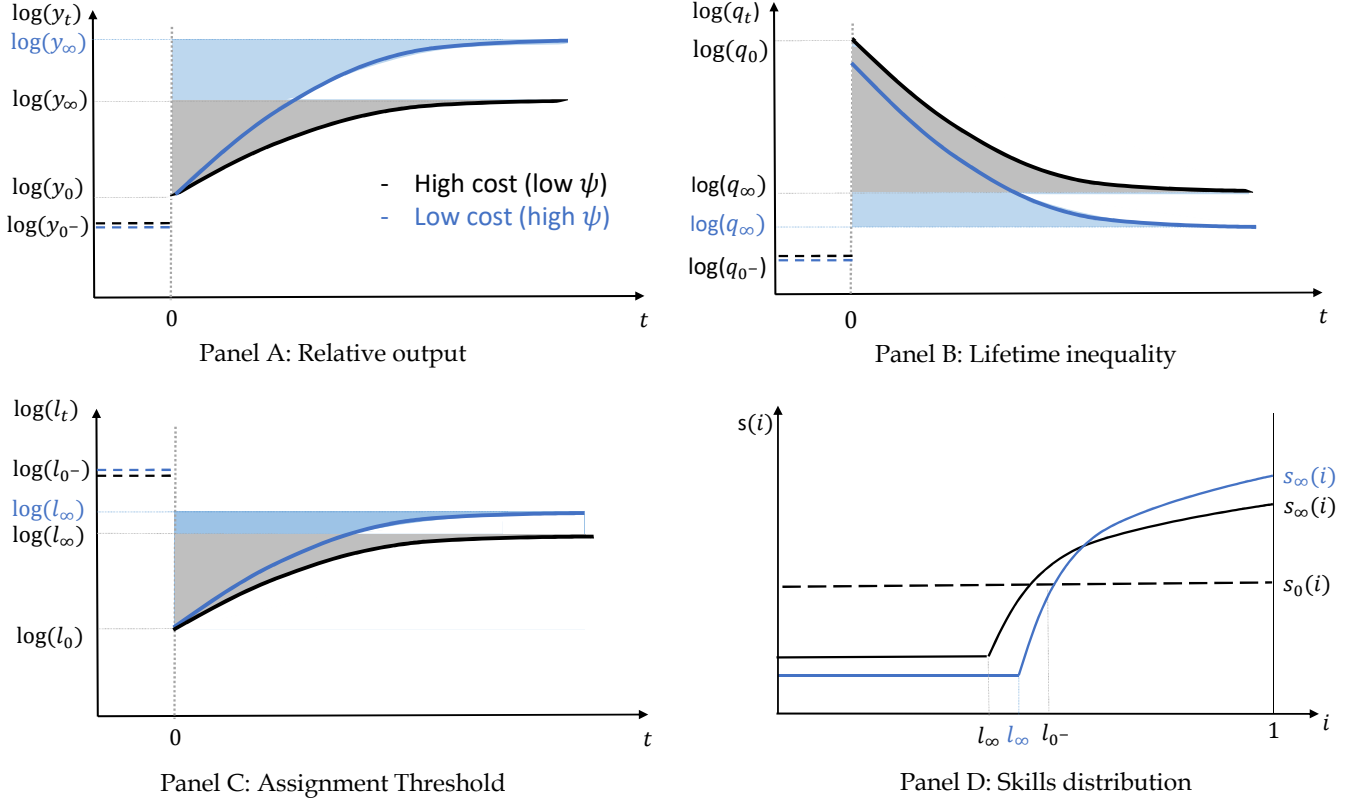
4.2 Comparative statics with respect to the cost of skill investment

We now consider how the parameter ψ affects the economy's adjustment to a skill-biased technological innovation. This comparative static exercise speaks to differences across historical episodes in the gap between younger and older generations' ability to invest in skills. Specifically, it captures situations in which younger generations may have found it easier to invest in skills in high demand than older generations due to, for example, better educational systems, the availability of vocational training for young workers, or the absence of re-training programs for older generations.

Figure 3 illustrates the responses of two economies that differ with respect to the skill investment cost of young generations. The blue lines depict the adjustment of an economy with a low investment cost (i.e., high value of ψ), and the black lines represent the responses of an economy with a high investment cost (i.e., low value of ψ). Appendix A.7 presents a formal proposition supporting the graphical representation in Figure 3.

In the short-run, both economies exhibit identical responses in relative output and worker allocation. This follows from the fact that ψ does not affect the self-selection decisions of generations born before the shock. However, a higher ψ attenuates the short-run increase in lifetime inequality because future relative wages fall by more due to the larger increase in the future supply of high-i skills implied by the more responsive skill lottery in (21). The larger change in the skill distribution of the economy with a lower investment cost (i.e., higher ψ) has two important implications for its dynamic adjustment to the shock. In the long run, it implies that relative output (lifetime inequality) increases more (less). Along the transition, it implies a larger cumulative impulse response in both lifetime inequality and relative output. Thus, as the next theorem shows, economies with a lower cost of skill investment for younger workers exhibit slower, more backloaded adjustment in relative output and inequality.

Figure 3: Comparative statics with respect to ψ

Panel A: Relative output

Panel B: Lifetime inequality

Panel C: Assignment Threshold

Panel D: Skills distribution

Theorem 2.2 (DCIR comparative statics with respect to ψ) *Following a skill-biased technological innovation $\Delta \log(A)$, lifetime inequality (q) and relative output (y) adjust slower in economies with a lower cost of skill investment for younger workers (i.e., higher ψ). Formally,*

$$\frac{\partial DCIR(q)}{\partial \psi} > 0, \quad \frac{\partial DCIR(y)}{\partial \psi} > 0$$

Proof. The proof is analogous to the one for Theorem 2.1 but using Proposition A.2 in Appendix A.7 instead. ∎

4.3 Back to LeChatelier Principle and the risks of extrapolation

To better understand the comparative statics with respect to η and ψ, it is useful to return to the supply-demand representation of the economy's adjustment introduced in Section 3.3. The different dynamic implications of changing η or ψ arise because the two parameters shape different horizons of the reduced-form elasticity of relative output supply. Figure 4 illustrate these implications.

Both higher values of η and ψ increase the elasticity of relative supply in the long-run. However, the timing of the increase in φ_t differs when the economy has a higher η or a higher ψ. Specifically, increasing η flattens the path of φ_t but increasing ψ steepens it. Intuitively, a higher η front-loads more the response in the relative supply of H goods by making it easier

Figure 4: Effect of η and ψ on the elasticity of relative output supply (φ_t)
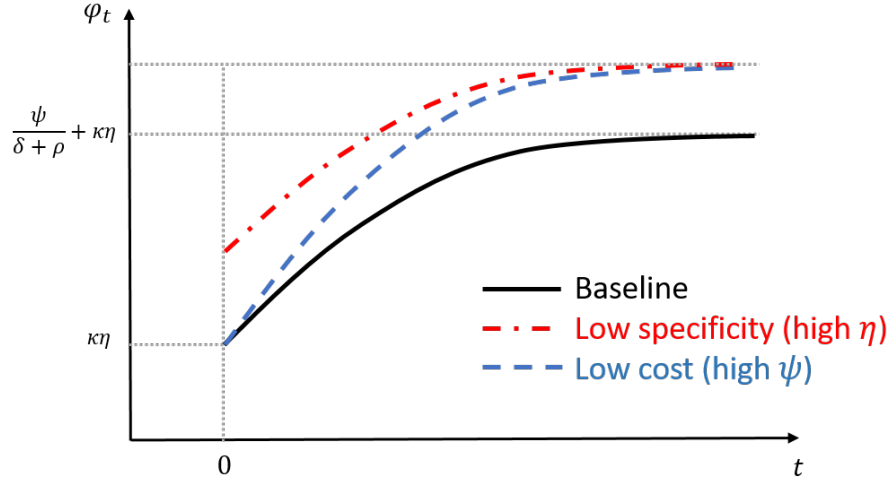
for skill types to reallocate across technologies in response to the shock. This in turn reduces the relative wage changes and, therefore, reduces the incentives of young workers to invest in skills. As a result, there are smaller differences in the skill distribution across generations and φ_t is more similar across horizons. In contrast, a higher ψ implies that it is easier for new generations to invest in skills, amplifying the difference in the skill distribution across generations and making φ_t more different across horizons.

This discussion also points to which type of economies should cause researchers to exercise more caution when extrapolating from observed responses at short horizons. This is the case precisely when technology-skill specificity is higher and/or the cost of skill investment for young generations is lower. In such economies, increases in relative output and decreases in relative wages are slower and more back-loaded.

4.4 Additional determinants of skill distribution dynamics

The theory so far has ignored several determinants of the dynamics of the skill distribution along the transition. In Appendix B, we consider three extensions that relax some of the assumptions of our baseline model. For all extensions, the results above regarding the economy's comparative statics with respect to η and ψ remain valid.

Our first extension considers a "learning-from-others" externality. Specifically, we relax the assumption that the reference distribution $\bar{s}_t(i)$ in the skill investment problem is exogenous and fixed over time. Instead, we assume that certain skills may be easier to acquire than others because workers "learn from others" when such skills are already abundant in the economy.[22] This extended model yields impulse response functions that are qualitatively similar to those of our baseline economy when ψ is higher and δ is lower. Thus, compared

---

[22]This mechanism implies that the skill distribution of existing workers directly affects the skill investment problem in (9) for incoming generations. Such effects could make more abundant skills more attractive when skills of different generations are complements in production, as in Chari and Hopenhayn (1991), or make them less attractive when skills of different generations are substitutes in production, as in Card and Lemieux (2001).

to our baseline model, learning from others makes the adjustment slower, with higher y_t and lower q_t in the long-run.

Our second extension relaxes the assumptions that workers can only invest in new skills upon birth by allowing a fraction of older generations to re-train when technology changes at $t = 0$. This extension yields impulse response functions that are qualitatively similar to those of our baseline economy when η is higher (i.e., lower technology-skill specificity).

Our third extension allows for population growth by making the birth rate higher than the death rate. When population growth is higher, the convergence rate λ is also higher, implying a faster adjustment for relative output and inequality, but no change in their long-run adjustment.

5   Observable Implications at Short Horizons: changes within- and between-generations

The previous sections have shown that the degree of technology-skill specificity and the skill investment cost of young generations determine the extent to which economies adjust to skill-biased technological innovations either through the reallocation of workers within a generation or changes in the skill distribution across generations. In this section, we ask two related questions. What are the observable implications of our theory that are informative about η and ψ? Given observed responses, how can we tell if the economy's adjustment will be slower and more back-loaded?

There are a number of practical challenges in addressing these questions. The first is that, typically, researchers can only credibly measure the effects of new technologies over short horizons. This is either because the new technologies are recent and the transition is ongoing, or because ex-post it is hard to separate such effects from other confounding shocks at long horizons. The second is that many of the theoretical objects in the previous sections are hard to measure in the data without strong assumptions. For example, the relative wage ω_t in efficiency units is different than the relative average labor income (which is measurable) because of "selection" forces.[23] Furthermore, as in the q-theory of capital investment, q_t is a forward-looking variable whose measurement requires knowledge of the entire equilibrium path of ω_t. So, to construct q_t, we would need to observe ω_t along the entire transition to the new stationary equilibrium. Finally, the direct measurement of the skill distribution $s_t(i)$ and the technology-skill assignment l_t requires taking an explicit stance on observable attributes that determine worker skills in different activities (e.g., college graduation or occupation history). The empirical analysis is misspecified whenever the chosen attributes do not completely determine the relative productivity of workers in the two technologies.

Given these challenges, we focus on observable changes in relative payroll and relative employment across sectors/occupations over short horizons. These variables are common in

---

[23]In fact, recent empirical applications of assignment models that use log-linear functions $\sigma(i)$ and $\alpha(i)$ yield identical distributions of labor income across technologies – e.g., see Hsieh et al. (2013) and Burstein, Morales, and Vogel (2016).

most datasets and available across many countries and time periods. Our novel measurement insight is that relative employment changes for different worker generations are connected to the degree of technology-skill specificity and the cost of skill investment. As a result, even though we observe these changes at short horizons, they are informative about how economies adjust at longer horizons.[24]

Specifically, we consider two types of responses in the short-run: (i) the within-generation change in outcomes of the "old" generation born before the shock, and (ii) the between-generation difference in outcomes of "young" generations entering at the time of the shock and "old" generations born before the shock. For relative employment in high-tech production, we formally define the within- and between-generation elasticities as

$$\varepsilon_0^{within} \equiv \frac{1}{\Delta \log(A)} \left(\log \left(\frac{\int_{l_0}^1 s_0(i)di}{\int_0^{l_0} s_0(i)di} \right) - \log \left(\frac{\int_{l_{0^-}}^1 s_0(i)di}{\int_0^{l_{0^-}} s_0(i)di} \right) \right),$$

$$\varepsilon_0^{between} \equiv \frac{1}{\Delta \log(A)} \left(\log \left(\frac{\int_{l_0}^1 \tilde{s}_0(i)di}{\int_0^{l_0} \tilde{s}_0(i)di} \right) - \log \left(\frac{\int_{l_0}^1 s_0(i)di}{\int_0^{l_0} s_0(i)di} \right) \right).$$

The following theorem shows how these measures are affected by changes in η and ψ.

Theorem 3 (Observable implications in the short-run)

1. *Within-generation elasticity*

$$\frac{\partial |\varepsilon_0^{within}|}{\partial \eta} > 0, \quad \frac{\partial |\varepsilon_0^{within}|}{\partial \psi} = 0$$

2. *Between-generation elasticity*

$$\frac{\partial |\varepsilon_0^{between}|}{\partial \eta} < 0, \quad \frac{\partial |\varepsilon_0^{between}|}{\partial \psi} > 0$$

Proof. See Appendix A.9. ■

The theorem states that, in the short-run, economies with a higher degree of technology-skill specificity (i.e., lower η) experience weaker within-generation growth in the relative employment of older workers, but stronger between-generation differences in the relative employment of younger and older workers. Intuitively, as discussed above, the higher technology-skill specificity diminishes the reallocation of skill types across technologies, which reduces the re-allocation of older workers, amplifies the short-run increase in relative wages,

---

[24]The estimation of generation-specific responses are common in empirical analysis of how economies adjust to different types of shocks – e.g., Kim and Topel (1995), Card and Lemieux (2001), Autor and Dorn (2009), Autor, Dorn, and Hanson (2013), Guvenen et al. (2017), McCaig and Pavcnik (2018), Greenland, Lopresti, and McHenry (2019), Porzio and Santangelo (2019). Our insight is not that such responses can be estimated, but that, through the lens of our theory, they are connected to structural parameters and are thus informative about the economy's dynamic adjustment to technological innovations.

and, consequently, increases skill differences across generations. In contrast, the lower investment cost (i.e., higher ψ) does not affect older workers whose skills were chosen before the shock, so within-generation employment change remains the same. The lower cost however triggers larger changes in the skill investment decisions of younger generations, creating stronger between-generation differences in relative employment.

We can now use these results to return to the questions in the beginning of this section. Suppose that we observe a technological transition with small within-generations employment elasticity, but large between-generation employment elasticity. The results above indicate that such observed responses are consistent with high technology-skill specificity and/or small costs of skill investment for young workers. As a consequence, through the lens of our theory, this technological transition should have a more unequal and slower adjustment. Alternatively, transitions with both relatively large within- and between-generation elasticities are consistent with low technology-skill specificity and small costs of skill investment, leading to a slower adjustment but only a small short-run increase in inequality.

Finally, note that the within- and between-generation employment elasticities are defined with respect to the shock, $\Delta \log(A)$. This is the relevant measure of the magnitude of the shock inducing relative employment responses in the economy. Yet, in some applications (like the one in the following sections), the shock is not directly observed. To circumvent this empirical challenge, we can use the observed response of relative payroll in Proposition 1 since it depends on the parameters of technology-skill specificity and skill investment costs, as well as the size of the shock.

6 Application: Cognitive-biased Technological Transitions

Our theoretical results established that technology-skill specificity and the cost of skill investment for entering generations are connected to within- and between-generation differences in relative employment following skill-biased technological innovations. In this section, we provide three pieces of evidence indicating the extent to which these two determinants affected the adjustment of developed economies to new cognitive-biased technologies.

First, in 18 developed countries, employment growth in the most cognitive-intensive occupations was stronger for young workers than for old workers. Second, turning to a detailed investigation of these responses in Germany, we show that in the cross-section of occupations, growth of employment and payroll was increasing in the time spent performing cognitive-intensive tasks. We find that these responses are stronger for younger than for older generations. In line with our theory's predicted changes in skill investment, we use the unique features of the large-scale German training system to document higher growth in the number of trainees in more cognitive intensive occupations. Finally, we explore cross-regional variation in adoption timing to estimate empirical impulse response functions to one cognitive-biased technological innovation: the arrival of broadband internet in the early 2000s. We find that the impact on relative employment is small for older generations at all horizons, but

increasing over time for younger generations. Taken together, the evidence suggest that, for recent cognitive-biased innovations, technology-skill specificity is high and the cost of skill investment is smaller for younger generations.

6.1 Cognitive-intensive employment growth in developed economies

We define cognitive-intensive occupations as being the set of production activities that were disproportionately augmented by recent technological innovations. In our theory, we will interpret these activities as those corresponding to high-tech production. Our approach follows an extensive literature documenting that the recent arrival of new technologies in the workplace, like the computer and the internet, had different effects on jobs with different task content —e.g. Autor, Levy, and Murnane (2003), Spitz-Oener (2006), Autor and Dorn (2013), Akerman, Gaarder, and Mogstad (2015), and, for a review, Acemoglu and Autor (2011). Specifically, this literature has documented that these new technologies augmented the productivity of cognitive-intensive jobs whose daily activities require problem-solving, creativity, or complex interpersonal interactions. On the other hand, these recent technological innovations substituted for 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.[25]

We analyze the evolution of the occupation employment composition of 18 developed countries. We use data on the number of males employed by occupation for two age groups: "Young" workers aged 15-39 yrs and "Old" workers aged 40-64 yrs. We consider employment in 9 aggregate occupation groups.[26] Using the German BERUFNET dataset, we rank occupations according to their share of time spent on tasks that intensively require analytical non-routine and interactive skills. We classify as cognitive-intensive the top 3 occupations in this ranking: Managers, Professionals, Technicians and Associate Professionals.[27]

Figure 5 displays the recent trends of employment in cognitive-intensive occupations for several developed countries. The dashed bars indicate that employment in cognitive-intensive occupations has been expanding in 16 out of the 18 countries in our sample. This trend is a reflection of the occupation polarization process documented by Goos, Manning, and Salomons (2009) for European countries, Autor and Dorn (2013) for the United States, and Green and Sand (2015) for Canada.

---

[25]In Appendix C.3, we use the German Qualification and Working Conditions Survey to show that internet and computer usage is strongly correlated with time spent on cognitive tasks across occupations. We also document that there are no systematic differences in internet and computer usage across different cohorts of workers employed in the same occupation.

[26]Our sample of countries includes Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, United Kingdom, United States. As data sources, we use Eurostat for European countries and IPUMS International for Non-European countries. For all countries, these data sources report the number of persons employed in the following 2-digit ISCO occupations: Managers, Professionals, Technicians and Associate Professionals, Clerical Workers, Service and Sales Workers, Skilled Agricultural Workers, Craft Trades workers, Plant and Machine Operators, and Elementary Occupations.

[27]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 define an occupation's cognitive intensity as the simple average of the time spent on analytical non-routine and interactive tasks in the years of 2011-2013.

Figure 5: Recent trends in cognitive-intensive employment growth in developed countries
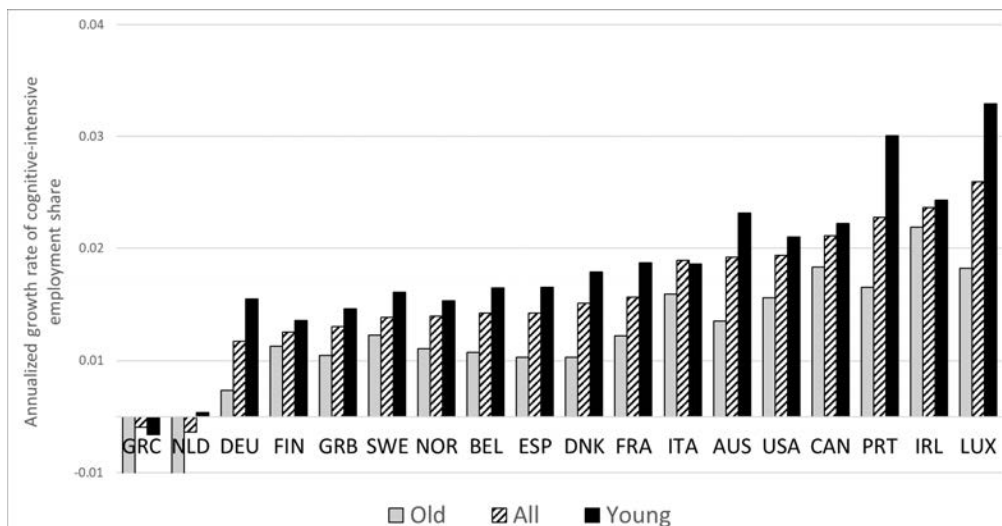
Note. The figure reports the log-change in the share of males employed in cognitive-intensive occupations in 1997-2017 for European countries, in 2000-2010 for the United States, and in 2001-2011 for Canada. Sample of males in two age groups: "Young" workers aged 15-39yrs and "Old" workers aged 40-64yrs. Cognitive-intensive occupations defined as the 3 occupation groups spending more time performing cognitive tasks on the job among the 9 occupation groups in the 2-digit ISCO classification: Managers, Professionals, Technicians and Associate Professionals. For each country, annualized growth rate is the log-change of the cognitive-intensive employment share in the period divided by the number of years.

Figure 5 also shows how annualized growth in the employment of cognitive-intensive occupations differed for younger and older generations of workers. While older workers increased their employment in cognitive-intensive occupations in most countries, this increase was substantially stronger for younger generations. Across all countries, the annualized growth in cognitive-intensive employment of younger workers was 73% higher than that of older workers. The young-old gap is higher whenever overall reallocation is higher: across countries, there is a correlation of 0.43 between the young-old gap in cognitive-intensive employment growth and that of all workers. These new stylized facts complement the finding in Autor and Dorn (2009) that the average age of workers employed in contracting middle-wage occupations increased in the United States between 1980 and 2005.

As discussed in Section 5, the different employment responses for young and old workers is consistent with an elastic supply of cognitive skills in the long-run driven by younger generations tilting their investment towards skills used in cognitive-intensive occupations. However, the aggregate trends in Figure 5 are subject to concerns about potential confounding shocks causing the expansion of cognitive-intensive employment. They also do not provide any direct evidence about the skill investment mechanism in our theory. Moreover, by not relying on a specific innovation, they are not informative about the dynamic adjustment of economies to new technologies. For these reasons, we now turn to a more detailed investigation of the impact of cognitive-biased technologies on the German labor market.

6.2 Cognitive-intensive employment growth and new technologies: Evidence from Germany

We next study how the German economy adjusted to recent cognitive-skill-biased technological shocks. We first describe the data used in our analysis. We then investigate the relative performance of occupations with a higher cognitive intensity in terms of employment, payroll, and numbers of trainees. Finally, we exploit quasi-experimental cross-regional variation in adoption timing of broadband internet to estimate the dynamic impact of this new technology on cognitive intensive occupations.

6.2.1 Data

Our main source of information on German labor market outcomes is the LIAB Longitudinal Model between 1995 and 2014. We follow Card, Heining, and Kline (2013) to construct a sample of employed males aged 15-64 years old residing in West Germany. We first use individual-level information to construct yearly series of outcomes for 120 occupations. While our theory features only two technologies, this is an abstraction, and we obtain more variation empirically by using more detailed occupation information. We therefore now move from the sharp predictions of the two-technology theory to look at employment trends across occupations more generally. Appendix C.1 lists the steps involved in constructing our sample.

We then construct a second dataset with annual data on employment and payroll for each occupation in 323 regional labor markets. Following Dauth, Findeisen, and Suedekum (2014) and Huber (2018), we use administrative districts to define regional labor markets in West Germany.[28] We use the BERUFNET dataset discussed previously to define each occupation's cognitive intensity as the share of time spent performing analytical non-routine and interactive skills.

We construct labor market outcomes for different groups of individuals in our sample. We first construct outcomes for "Young" and "Old" generations of workers. We define the "Young" generation as all individuals born after 1960, and the "Old" generation as all other individuals in the sample. The young generation was at most 35 years old in the beginning of our period of analysis in 1995, representing 57.5% of the German labor force in that year. Over time, the young generation increased its overall employment share, reaching around 89% by the end of the analysis period in 2014 (when the young generation was at most 54 years old). Appendix C shows that the results in this section are robust to alternative definitions of worker generations.

We also define a sub-sample of trainees composed of workers whose employment status is trainee, student trainee, or intern. In this trainee sub-sample, 97.5% of all workers are below 30 years old and the mean age is 20.8 years old. Thus, we interpret the changes in occupation

---

[28]We construct our data using the district of the establishment of the main job of each individual in any given year. Since this information is only available after 1999, we use the establishment's district in 1999 to construct the worker's district affiliation in 1995-1998.

allocation in this trainee sample as a proxy for the changes in the skill investment decision of incoming generations in our theory.[29]

6.2.2 Cognitive intensity and labor market outcomes across occupations

We now study the relationship between employment growth and cognitive intensity across occupations in Germany. Motivated by the predictions in Section 5, we estimate the following linear regression for each worker generation g and year t:

$$\log Y_{o,t}^g - \log Y_{o,1995}^g = \beta_t^g \bar{C}_o + \epsilon_{o,t}^g \tag{28}$$

where $Y_{o,t}^g$ is a labor market outcome in occupation o at year t of workers of generation g, and $\bar{C}_o$ is the cognitive intensity of occupation o.[30]

Table 1 reports the estimation of equation (28) in the periods of 1995-2000 (Panel A) and 1995-2014 (Panel B). We report the estimated impact of the occupation's cognitive intensity on the log-change of employment in columns (1)–(3), payroll in columns (4)-(6), and number of trainees in column (7).

Over all horizons, column (1) indicates that occupations with a higher cognitive intensity experienced stronger growth in employment. This differential employment response is larger over longer horizons. Compared to the least cognitive-intensive occupation, the employment growth in the most cognitive-intensive occupation was around 143 percent higher by 2014. These results show that the German trends in Figure 5 also hold when we consider variation across occupations with different levels of cognitive intensity. Comparing the responses by generation in columns (2)–(3), we find that employment growth in cognitive-intensive occupations was weaker for older generations than for younger generations. In fact, the coefficient estimates for the old generations are between one-third and one-half of that for the young generations at all horizons. Column (2) also shows that young generations display very strong employment growth in occupations with a higher cognitive intensity in all sample periods.

Columns (4)-(6) show that the relative payroll responses are slightly stronger than the relative employment responses between 1995 and 2014. This suggests that there were only small relative changes in the average earnings of those employed in cognitive-intensive occupations. As discussed in Section 2.1, in our theory, these relative payroll responses include the rise in the marginal productivity of labor in more cognitive-intensive occupations, as well as the change in overall productivity of workers employed in cognitive-intensive occupation (i.e., the selection effect created by the change in worker allocations). So, the difference between columns (4) and (1) do not correspond to the response of the relative wage per efficiency unit of more cognitive-intensive occupations. In fact, the small responses in relative average earn-

---

[29]In line with this interpretation, Eckardt (2019) shows that the occupation classification of trainees in Germany is unique in its ability to capture a worker's field of study.

[30]We do not include any controls in our baseline specification. Appendix Table B4 shows that results are similar when we include controls that capture potential confounding effects from the occupation's exposure to immigration and trade shocks in the period of analysis.

Table 1: Cognitive intensity and labor market outcomes across occupations in Germany

Dependent variable:	Employment Growth			Real Payroll Growth			Trainee
	All	Young	Old	All	Young	Old	Growth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Change in 1995-2000							
Cognitive intensity	0.388***	0.650***	0.113***	0.340***	0.616***	0.157***	0.379*
	(0.076)	(0.098)	(0.043)	(0.048)	(0.070)	(0.037)	(0.209)
Panel B: Change in 1995-2014							
Cognitive intensity	1.488***	1.894***	0.871***	1.535***	2.029***	1.044***	2.121***
	(0.225)	(0.234)	(0.229)	(0.227)	(0.238)	(0.223)	(0.385)

Note. Sample of 120 occupations. Each panel reports the estimate for the dependent variable over the indicated time period. Young cohort defined as all workers born after 1960 and Old cohort as all workers born before 1960. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

ings for both young and old are consistent with strong selection forces created by entry of marginal workers with lower occupation-specific productivity than infra-marginal workers. Such a pattern arises in assignment models with a Frechet distribution of technology-specific ability, as in Hsieh et al. (2013) and Burstein, Morales, and Vogel (2016).

Column (7) shows that occupations with a higher cognitive intensity experienced stronger growth in the number of trainees. As discussed above, trainee programs are an important part of the formal training of young individuals in Germany – especially for non-degree occupations. As such, our estimates suggest that investment in cognitive-intensive skills by incoming generations became stronger throughout this period.

Taken together, this evidence again speaks qualitatively to the main mechanisms in our model. The small responses in employment for old workers suggest that skills are very specific to occupations with the same cognitive content. The large differences in employment responses between generations suggest that the cost of skill investment is smaller for younger workers. In fact, young workers seem to increase their investment on cognitive skills by becoming trainees in occupations with a higher cognitive intensity. The differences in overall responses at longer horizons are in line with LeChatelier's principle.

That said, one concern with this interpretation is that the occupation-level responses may not be a consequence of a single technological innovation in a particular year. Instead, they may be driven by different innovations introduced sequentially throughout the period of analysis – e.g. computers, industrial robots, or the internet. Thus, while our interpretation remains qualitatively valid, it is hard to quantitatively connect the estimates above to the mechanism in our theory because the empirical estimates are not impulse response functions to one-time permanent shocks. That is, the estimated dynamics may potentially confound both the endogenous skill distribution dynamics and the exogenous sequence of technological innovations. We address this concern in the next section.

Additional results. Appendix C.4 presents additional results that attest the robustness of the findings presented in this section. Table B2 shows that results are qualitatively similar over different horizons. Second, Tables B3 shows that the positive relative employment growth in cognitive occupations is driven by the top third of occupations by cognitive intensity (for all workers and separately for each worker generation). Table B4 shows that results are similar when restricting the sample to native-born Germans, changing the definition of the young generation, or including controls for each occupation's exposure to trade and immigration shocks.

6.2.3 Dynamic adjustment to broadband internet adoption

In this section, we analyze the dynamic response to one cognitive-biased technological innovation: the introduction of broadband internet in the early 2000s.[31] There are two main reasons to focus on this particular innovation in Germany. First, it resembles the one-time permanent shock studied in Section 3 since its adoption was fast: the share of households with broadband access increased from 0% in 2000 to over 90% in 2009. Second, it is possible to explore cross-regional variation in adoption timing to estimate the impulse response functions of labor market outcomes for different worker generations. Our strategy relies on the fact that the timing of broadband adoption 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%. In addition, following Falck, Gold, and Heblich (2014), we isolate exogenous spatial variation in adoption timing implied by the suitability of pre-existing local telephone networks for broadband internet transmission.

Empirical Strategy. Our goal is to estimate the dynamic impact of broadband internet adoption on labor market outcomes across districts in Germany. For each year between 1996 and 2014, we estimate the following linear specification

$$Y_{io,t}^g - Y_{io,1999}^g = \sum_{c \in \{\text{young, old}\}} (\alpha_t^c + \beta_t^c \bar{C}_o) 1_{[g=c]} DSL_i + \delta_{o,t} + \zeta_{g,t} + X_{io,t}^g \gamma_t^g + \epsilon_{io,t}^g, \quad (29)$$

where o is an occupation, i is a German district, and g is a worker generation. In this specification, $Y_{io,t}^g$ is a labor market outcome (employment or payroll), DSL_i is the broadband internet penetration in district i in 2005 (normalized to have standard deviation of one), and $\bar{C}_o$ is the time-invariant measure of the cognitive intensity of occupation o. As above, we consider two generations: the old generation born before 1960 and the young generation born after 1960. The specification includes two sets of fixed-effects: generation-year fixed effects that capture nationwide labor market trends of different worker cohorts, and occupation-year fixed-effect that absorbs any confounding shock that has the same impact on an occupation

---

[31] As shown Akerman, Gaarder, and Mogstad (2015), broadband internet expanded the relative demand for more educated workers in non-routine jobs inside firms. In Appendix C.3, we show that this new technology is disproportionately used by individuals employed in more cognitive-intensive occupations.

in all regions. We also include a control vector $X_{io,t}^g$ to absorb confouding effects associated with the pretrend growth in 1995-1999 and initial district demographic characteristics. These controls account for differential performance of cognitive-intensive occupations in regions with characteristics that may affect the profitability of broadband internet adoption.[32]

We are mainly interested on the impact of broadband internet adoption on the relative outcome of cognitive-intensive occupations for each generation: β_t^g in equation (29). 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^g is the difference between regions A and B in the relative outcome of workers of generation g in more cognitive intensive occupation.

The unbiased estimation of β_t^g requires an exogenous source of variation on the adoption of broadband internet across German districts in 2005. However, 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 a MDF located more than 4200m away. Thus, areas initially located close to MDFs obtained broadband internet access before areas located far away 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.[33] Building on this idea, we construct two instrumental variables at the district-level that measure the

---

[32]We follow Dix-Carneiro and Kovak (2017) and Freyaldenhoven, Hansen, and Shapiro (2018) by explicitly controlling for pretrends. As argued by the latter paper, pretrends caused by unobserved confounding effects might exist even when they are not actually observed in the data due to estimation error, implying they should be controlled for in estimation. 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 B7 shows results for different control sets.

[33]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.

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) 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."

We then estimate (29) using the exogenous variation induced by these two measures of the cost of expanding broadband availability in the district, which we summarize in the vector Z_i. Specifically, since the observation unit in equation (29) is an occupation-generation-district triple, we use an instrument vector that includes Z_i interacted with generation dummies and the cognitive intensity of each occupation o, $\bar{C}_o$. Figure B2 in Appendix C.1 presents the pattern of cross-district variation in Z_i. Table B5 in Appendix C.4 shows that regions with higher values of these cost measures had a lower share of households with broadband access in 2005. Intuitively, this is the main source variation in the first-stage of our strategy to estimate equation (29). Formally, to test for weak instruments for the multiple endogenous variables in (29), Table B6 in Appendix C.4 shows that we obtain high values for the Sanderson-Windmeijer F-statistics (Sanderson and Windmeijer, 2016).

Results. We now turn to the estimation of β_t^{old} and β_t^{young} in equation (29). Panel A of Figure 6 reports the estimates for each year between 1996 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, there is a significant impact on the relative employment of young cohorts 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 0.5 log-points more young workers employed in the most cognitive-intensive occupation than in the least cognitive-intensive occupation. However, we do not find such an effect for old cohorts – if anything, the effect is negative. Thus, our estimates indicate a positive between-generation difference in relative employment growth. Figure B7 in Appendix C.4 shows that the between-generation difference is statistically significant for every year after 2006.

We can use the theoretical predictions in Section 5 to interpret these empirical results. The small relative employment response of old generations suggests that technology-skill specificity is high (i.e., η is low). In this case, most workers from old generations do not switch occupations as their skills would have a lower value in the more cognitive-intensive

occupations augmented by the technological innovation. Furthermore, the positive between-generation difference in the relative employment response indicates that incoming cohorts tilt their investments towards skills more suitable for cognitive-intensive jobs. This suggests that the technological innovation induced skill heterogeneity across generations (i.e., ψ is positive).

In Panel B of Figure 6, we investigate how early broadband expansion affected the relative payroll of more cognitive intensive occupations for *all* worker generations. Specifically, we estimate equation (29) with a single generation c containing workers of all ages in the district. This is the empirical analog of the impulse response function for relative output y_t presented in Section 3. Again, we find no evidence of responses in the pre-shock period of 1996-2005. Starting in 2006, there is a slow and steady increase in the relative payroll of more cognitive-intensive occupations.[34] In our theory, these results are consistent with broadband internet augmenting the relative productivity of cognitive intensive occupations when cognitive and non-cognitive intensive occupations are substitutes in production (i.e., $\theta > 1$).

Figure 6: Impact of early DSL adoption on more cognitive-intensive occupations
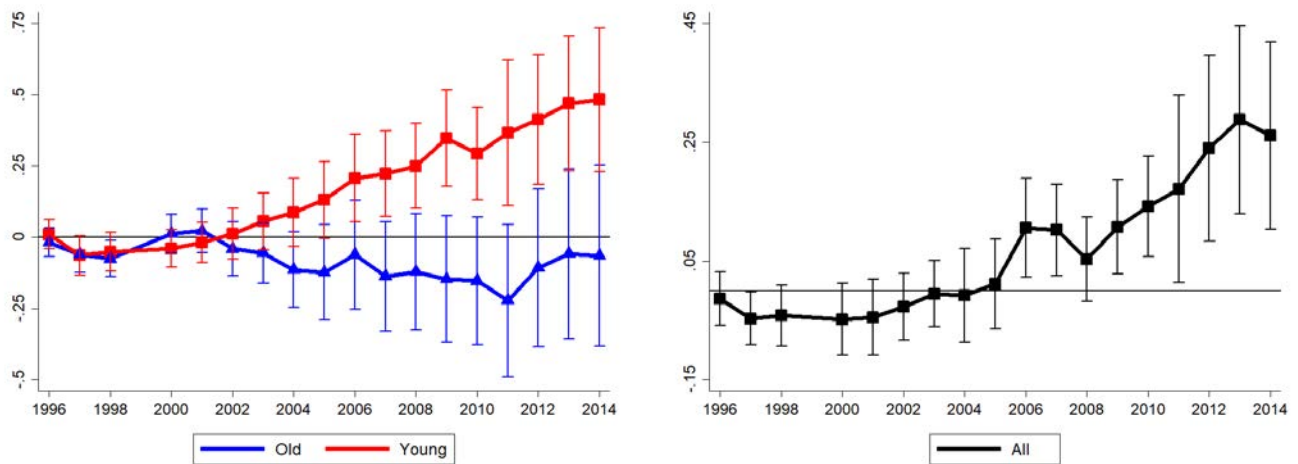
(a) Relative employment response for each generation   (b) Relative payroll response for all generations

Note. Left panel: estimation of equation (29) for log-employment as dependent variable in the sample of 2 generations, 120 occupations and 323 districts. Right panel: estimation of equation (29) for log-payroll as dependent variable in the sample of 120 occupations, 323 districts, and a single generation with all working-age employed individuals. For each year, the dot is the point estimate of β_t^g. All regressions are weighted by the district population size in 1999 and include occupation-time and generation-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. Bars are the 90% confidence interval implied by the standard error clustered at the district level.

Additional results. Appendix Tables B7–B9 provide additional results that complement the estimates in Figure 6. Table B7 shows that results are qualitatively similar when we drop the pretrend control, but estimated coefficients are less precise and slightly smaller in magnitude. We also demonstrate that results are similar when controlling for district-generation-year fixed-effects. This is reassuring as this restrictive set of controls absorbs all potential con-

---

[34]Figure B8 in Appendix C.4 shows that we obtain similar qualitative responses for the relative employment of more cognitive intensive occupations among all generations of workers in the district.

founding shocks that affect each district-generation pair in a year. In this case, identification comes purely from the differential effect of early broadband expansion on occupations with a higher cognitive intensity. That is, this control set captures any variation that might have resulted in a district receiving broadband access early, including differential immigration into a district that received DSL or differential aging or birth patterns in the district over time.

Table B8 presents results when we vary the definition of the young generation, as well as when we restrict the sample to only native-born males. We consider several definitions of the young generation: those born after 1955, 1965 or 1970, and those aged less than 35, 40 or 45 in each year. Once again, the results are qualitatively similar across specifications. However, in line with our theory, the estimated coefficients for the young generation are stronger when we restrict the young generation to include only more recent cohorts.

In addition, Table B8 also reports estimates with outcomes in a sample that includes only workers employed in small establishments. The rationale is that larger establishments might have received broadband internet access prior to the roll-out for households across Germany. In this case, we would expect adjustment in these establishments to have occurred earlier, biasing our results to zero. Consistent with this intuition, we find that our results are quantitatively stronger in the sample of small establishments.

Finally, Table B9 shows that early adopting regions experienced stronger growth in the number of trainees in more cognitive intensive occupations. This evidence is consistent with our theory's prediction that, after the arrival of a cognitive-intensive innovation, incoming cohorts increase their investment on cognitive skills.

6.3 Is this time different?

The evidence above suggests that recent cognitive-biased innovations triggered a transition that is particularly slow and unequal. However, our theoretical results indicate that not all technological transitions are the same. The adjustment may be less unequal and faster if the economy features lower technology-skill specificity. The results in Section 5 indicate that, in this case, old and young generations exhibit more similar changes in relative employment across occupations along the transition. We now build on this insight to investigate whether past changes in employment composition featured weaker between-generation differences and, consequently, may have been part of a transition with lower technology-skill specificity.

We focus on the evolution of the occupation employment composition in Germany and the United States over different time periods. We again use the nine aggregate occupations in the 2-digit ISCO classification used in Section 6.1. Due to data availability, our early period is 1970-1987 for Germany and 1960-1990 for the United States. For both countries, the recent period is the same as the one used to compute the trends in Figure 5.[35] To obtain a

---

[35]Our primary data source for the early period is the individual-level Census data downloaded from IPUMS international, which contains information on the 2-digit ISCO occupation of males aged 16-64 in each Census year. For the recent period, we compute all outcomes using the same underlying data of Figure 5. We select a sample of employed males in each country-year and split them into two age groups: "Young" workers aged 15-39yrs and "Old" workers aged 40-64yrs.

measure of the expanding occupations that is consistent over time, we no longer rely on the set of cognitive intensive occupations since past shocks may have augmented a different set of skills. Instead, for each country and period, we define the expanding occupations as the three occupations with the highest change in log employment share among young workers. Through the lens of our theory, since young workers adjust their skills to work on occupations that became more attractive, their employment decisions provide a revealed-preference way of recovering the occupations experiencing positive demand shocks under the assumption of no shocks to the cost of investing on different skills.

Table 2 reports the employment growth trends in the three occupations with the highest growth among young workers. Columns (1) and (3) report substantial growth in these occupations for both periods and countries. Interestingly, columns (2) and (4) show that the two periods differ in the relative magnitude of the between- and within-generation components of employment changes. As in Figure 5, there is a large between-generation difference in recent years when most expanding occupations were cognitive intensive. However, such a between-generation difference was much smaller before 1990 when changes in the occupation composition was more similar for young and old generations.[36] In this earlier period, the set of expanding occupations was less cognitive intensive with services and retail occupations at the top of the list in both countries. In fact, Germany did not have any cognitive intensive occupation among the fastest growing occupations in 1970-1987.

Through the lens of our theory, these aggregate trends are consistent with a lower degree of technology-skill specificity in the occupations expanding before 1990. In this case, changes in the skill distribution across generations are smaller, giving rise to a faster and less unequal transition. It is important to notice that this is just one of many possible interpretations of the evidence in Table 2. For example, the generation-specific shocks in college graduation rates documented in Card and Lemieux (2001) may help explain why young and old generations have similar changes in employment composition in the 1970s and 1980s.

7   A Numerical Illustration of the Theory

We conclude the paper by using the evidence in the preceding section to analyze how economies adjust to cognitive-biased innovations. Our goal is not to provide a full quantitative account of such technological transitions, but rather to numerically illustrate our theoretical insights. In particular, we are interested in giving a sense of how large are the impacts of technology-skill specificity and skill investment cost on the economy's dynamic adjustment following technological shocks. In addition, by presenting the full non-linear equilibrium dynamics, the numerical exercise also demonstrates that our theoretical insights are not driven by the first-order approximations.

---

[36]Appendix C.5 provides additional evidence of this trend reversal by investigating the correlation between the change in the average age and the employment share across occupations in the United States over different time periods. We show that such a correlation is strongly negative in periods after 1990, but it was much weaker between 1960 and 1990.

Table 2: Changes in between-generation employment differences and employment shares across occupations in different periods

	Early period		Recent period	
	$\frac{1}{T}\Delta \log e_t^{all}$	$\frac{\Delta \log(e_t^{young}/e_t^{old})}{\Delta \log e_t^{old}}$	$\frac{1}{T}\Delta \log e_t^{all}$	$\frac{\Delta \log(e_t^{young}/e_t^{old})}{\Delta \log e_t^{old}}$
	(1)	(2)	(3)	(4)
Germany	1.94%	0.173	1.59%	0.456
United States	0.91%	0.206	1.45%	0.476

Note. Columns (1) and (3) report the annualized growth rate in the three 2-digit ISCO occupations with the highest change in log employment share in the period among young workers in the country (where T is the number of years in the period). For the top 3 occupations by log-employment growth for young workers, columns (2) and (4) report the ratio between the log-change in the between-generation employment share and the log-change in employment share for old workers. Early period: 1970-1987 for West Germany and 1960-1990 for the United States. Recent period: 1997-2017 for Germany and 2000-2015 for the United States. Sample of males in two age groups: "Young" workers aged 15-39yrs and "Old" workers aged 40-64yrs.

We map the H technology to cognitive-intensive occupations, and use the empirical impulse responses of Section 6 to parameterize the model. We first externally calibrate the discount rate ρ to match an annual interest rate of 2% and the demand elasticity of substitution to $\theta = 3$. We then select the parameters governing production technologies $(\alpha(i), \sigma(i))$ and the skill distribution dynamics (δ, ψ) to match the estimates in Figure 6. The decline in the share of the old workers in total employment from 1997 to 2014 implies $\delta = 0.057$, i.e, an expected working life-span of about 18 years after age 35. The small response in the cognitive-intensive employment of old generations yields an η close to zero, and the large young-old gap in the relative employment response implies $\psi = 0.35$. Appendix D describes the matching procedure in detail, along with the model's goodness of fit.

We use the parameterized model to study the consequences of a cognitive-biased innovation that increases the employment share in the cognitive-intensive technology from 20% to 50%.[37] We focus here on the impact of the shock on average welfare ($\Delta \bar{U}$) and lifetime welfare inequality ($\Delta \bar{\Omega}$), as well as the importance of transitional dynamics as measured by $DCIR(q)$.[38] Appendix D.3 shows the dynamics of the skill distribution and other outcomes.

Table 3 shows that, for our baseline parameterization, the increase in average welfare across all generations (in consumption equivalent units) is 46% and the increase in lifetime welfare inequality is 39%. These large effects follow from the substantial shock size necessary to induce the reallocation of almost one-third of the economy's labor force.

The remaining rows of Table 3 compare these figures to those obtained with two calculations that ignore the adjustment across generations. The 'Short-run' calculation assumes that changes observed at impact are permanent, while the 'Long-run' calculation assumes that the

---

[37]These values approximately correspond to the cognitive-intensive employment share in 1997 of the countries with the lowest and the highest cognitive-intensive employment share among those listed in Figure 5 (Portugal and Netherlands, respectively). Thus, our quantitative results can be seen as analyzing the transitional dynamics of a cognitive-biased shock that generates convergence in cognitive-intensive employment shares across such countries.

[38]Our analysis specifies the discount rate of social welfare to $r = \rho + \delta$, so that the social discounting of future generations is identical to the discounting of worker's future utility.

Table 3: Changes in Average Welfare and Lifetime Welfare Inequality

| | Baseline $(\eta \approx 0, \psi = 0.35)$ | | Low specificity $(\eta = 0.75, \psi = 0.35)$ | |
	$\Delta \bar{U}$	$\Delta \bar{\Omega}$	$\Delta \bar{U}$	$\Delta \bar{\Omega}$
True	46%	39%	44%	29%
Short-run	31%	76%	40%	45%
Long-run	55%	30%	47%	24%
$DCIR(q)$	0.9		0.4	

Note. The table reports the the changes in average welfare $\Delta\bar{U}$ and lifetime welfare inequality $\Delta\bar{\Omega}$ implied by a shock calibrated to increase the employment share in cognitive-intensive occupations from 20% to 50% between stationary equilibria. 'True' corresponds to the measures that fully account for the economy's transitional dynamics. 'Short-run' assumes that changes at impact are permanent. 'Long-run' assumes that long-run changes happened at impact.

changes observed in the long-run were permanent and happened at impact. As discussed in Section 3.3, these calculations are equivalent to those that would be obtained by researchers using a reduced-form model that ignores changes in the supply elasticity over time.

We can see that these two calculations lead to substantial biases in welfare analysis. The 'Short-run' calculation severely understates the average welfare gains and overstates the inequality increases. The opposite is true for the 'Long-run' calculation. The biases arise because of the slow adjustment in the economy's skill distribution. For instance, the $DCIR(q)$ of 0.9 reported in the last row implies that a worker born right before the shock expects to experience in her lifetime a relative wage that is 90 percent larger than the relative wage in the long-run equilibrium of the economy. Thus, the 'Short-run' approach misses the future accumulation of skills that increases relative output– thus reducing the ideal price index and increasing average real wages– and reduces relative wage of cognitive-intensive occupations. In contrast, the 'Long-run' approach misses the fact that it takes generations for the economy to accumulate the cognitive skills necessary to achieve the levels of relative output and wages observed in the long-run.

In the remaining columns of Table 3, we analyze the same shock in an economy with a lower degree of technology-skill specificity (i.e., higher η). As discussed in Section 5, in this case, the between-generation difference in the relative employment response is smaller due to the smaller change in the skill distribution across generations. As such, we interpret the comparison between our baseline and this alternative calibration as a numerical illustration of the welfare consequences of the same shock if a lower technology-skill specificity resulted in more similar occupation composition changes between old and young workers– as those reported in Section 6.3 for the United States and Germany before 1990. The second panel of Table 3 shows that the higher η implies a faster transition with $DCIR(q)$ falling from 0.9 to 0.4. This is a consequence of the stronger reallocation of old workers at impact which then leads to weaker increases in inequality and changes in the skill distribution, as well as

a stronger decline in the price index in the short-run (which translates into higher average welfare gain in the short-run). Importantly, the faster transition implies that there are smaller biases from the short- and long-run welfare calculations.

8   Conclusions

We develop a theory where overlapping generations of workers are heterogeneous over a continuum of technology-specific skills. Technological transitions are driven both by the reallocation of workers within a generation and changes in the skill distribution across generations. 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 skill-biased innovation, as well as derive observable predictions for changes in labor market outcomes within and between generations. We use these insights to study the adjustment of developed economies to recent cognitive-biased technological innovations. Several pieces of evidence show strong responses of cognitive-intensive employment for young but not old generations.

Taken together, we derive two broad takeaways from this piece. First, the evidence suggests that cognitive-biased transitions may be particularly unequal and slow to play out because of the high specificity of cognitive skills. Most of the adjustment happens through slow changes in the skill distribution across generations as opposed to the fast reallocation of workers within a generation. These features are not universal though. They may be different in past or future technological transitions where a broader set of skills can be transferred to the occupations or sectors improved by the technological innovation. Second, caution should be exercised when interpreting technological transitions based on evidence spanning much less than a generation. This may lead to overly pessimistic views of the consequences of new technologies for inequality and average welfare. Yet, observed changes for different generations, even at short horizons, are useful when combined with a theory of technological transitions. Looking at the decisions of younger workers allows us to "see the future" and thus appropriately derive the full implications of technological innovations.

Some of the ideas we presented may be useful to tackle other questions where transitional dynamics are important. Extensions of our theory with multiple sectors and regions can be used to study how various sources of labor demand shocks (e.g., trade liberalizations, routine-biased innovations) affect the economy over different horizons through within-generation worker reallocation and between-generation changes in skills. Moreover, future work can address normative questions related to the optimal speed of adjustment to such labor demand shocks or the role of workers with more transferable skills in providing aggregate insurance against them. Finally, the notion that within- and between-cohort changes at short horizons are informative about structural parameters governing elasticities at longer horizons can be used both to improve on empirical projections about future labor market conditions as well as to discipline other dynamic models with, for instance, incumbents and entrant firms.

References

Abramitzky, Ran, Victor Lavy, and Maayan Segev. 2019. "The Effect of Changes in the Skill Premium on College Degree Attainment and the Choice of Major." Tech. rep., National Bureau of Economic Research.

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.

Altonji, Joseph G, Peter Arcidiacono, and Arnaud Maurel. 2016. "The analysis of field choice in college and graduate school: Determinants and wage effects." In *Handbook of the Economics of Education*, vol. 5. Elsevier, 305–396.

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., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103 (6):2121–68.

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.

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.

Buera, Francisco J, Joseph P Kaboski, and Yongseok Shin. 2011. "Finance and development: A tale of two sectors." *American economic review* 101 (5):1964–2002.

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

Calvo, Guillermo A and Maurice Obstfeld. 1988. "Optimal time-consistent fiscal policy with finite lifetimes." *Econometrica: Journal of the Econometric Society* :411–432.

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. 2018. "Housing Booms and Busts, Labor Market Opportunities, and College Attendance." *American Economic Review* 108 (10):2947–94.

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. 2014. "Trade liberalization and labor market dynamics." *Econometrica* 82 (3):825–885.

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." .

Dvorkin, Maximiliano and Alexander Monge-Naranjo. 2019. "Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations." Tech. rep.

Eckardt, Dita. 2019. "Are Chemists Good Bankers? Returns to the Match between Training and Occupation." Tech. rep., LSE.

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.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2016. "The dynamics of inequality." *Econometrica* 84 (6):2071–2111.

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

Ghose, Devaki. 2019. "Trade, Internal Migration, and Human Capital: Who Gains from India's IT Boom?" .

Goldin, Claudia D. and Lawrence F. Katz. 2009. "The Future of Inequality: The Other Reason Education Matters So Much." Scholarly articles, Harvard University Department of Economics.

Goos, Maarten, Alan Manning, and Anna Salomons. 2009. "Job polarization in Europe." *American economic review* 99 (2):58–63.

Green, David A and Benjamin M Sand. 2015. "Has the Canadian labour market polarized?" *Canadian Journal of Economics/Revue canadienne d'économique* 48 (2):612–646.

Greenland, Andrew, John Lopresti, and Peter McHenry. 2019. "Import competition and internal migration." *Review of Economics and Statistics* 101 (1):44–59.

Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner. 2017. "Lifetime incomes in the United States over six decades." Tech. rep., National Bureau of Economic Research.

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.

Heckman, James J and Bo E Honore. 1990. "The empirical content of the Roy model." *Econometrica* 58 (5):1121–1149.

Heckman, James J, Lance Lochner, and Christopher Taber. 1998. "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents." *Review of economic dynamics* 1 (1):1–58.

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.

Itskhoki, Oleg and Benjamin Moll. 2019. "Optimal Development Policies With Financial Frictions." *Econometrica* 87 (1):139–173.

Kambourov, Gueorgui and Iourii Manovskii. 2009. "Occupational mobility and wage inequality." *The Review of Economic Studies* 76 (2):731–759.

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-II 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.

Lucas, Robert E and Benjamin Moll. 2014. "Knowledge growth and the allocation of time." *Journal of Political Economy* 122 (1):1–51.

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

McCaig, Brian and Nina Pavcnik. 2018. "Export markets and labor allocation in a low-income country." *American Economic Review* 108 (7):1899–1941.

McFadden, Daniel et al. 1973. "Conditional logit analysis of qualitative choice behavior." .

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

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

Perla, Jesse and Christopher Tonetti. 2014. "Equilibrium imitation and growth." *Journal of Political Economy* 122 (1):52–76.

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.