

Fast and Slow Technological Transitions*

Rodrigo Adão[†]
Booth School of Business

Martin Beraja[‡]
MIT

Nitya Pandalai-Nayar[§]
University of Texas at Austin

September 16, 2022

Abstract

Do economies adjust slowly to certain technological innovations and more rapidly to others? We argue that the adjustment is slower when innovations mainly benefit production activities requiring skills which are more different from those used in the rest of the economy. The reason is that, when such skill specificity is stronger, the adjustment of labor markets is driven less by the fast reallocation of older incumbent workers and more by the gradual entry of younger generations to the benefitted activities. We begin by documenting that the U.S. labor market adjusted differently to the arrival of Information & Communications Technologies (ICT) in the late twentieth century than it did to innovations in manufacturing at the beginning of that century. We then build an overlapping generations model of technological transitions. It allows us to sharply characterize the effects of skill specificity on equilibrium dynamics, match the evidence in a parsimonious way, and study its welfare implications. We show that stronger skill specificity helps to explain why the ICT transition was more unequal and slower, driven entirely by the gradual entry of younger generations who accrued more of the welfare gains from ICT innovations.

*We are grateful to Ha Bui and Agustin Gutierrez for excellent research assistance. For their valuable suggestions, we thank Daron Acemoglu, David Autor, Fernando Alvarez, Ariel Burstein, Juliette Caminade, V.V. Chari, Arnaud Costinot, Olivier Coibion, Steve Davis, Dave Donaldson, Oded Galor, Erik Hurst, Greg Kaplan, Larry Katz, Pablo Kurlat, Simon Mongey, Fabrizio Perri, Tommaso Porzio, Aysegul Sahin, Dejanir Silva, Christopher Tonetti, Venky Venkateswaran, Gianluca Violante, and Fabrizio Zilibotti.

[†]Email: rodrigo.adao@chicagobooth.edu. Web: www.rodrigoadao.com

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

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

1 Introduction

Technological transitions follow the arrival of major innovations. In the past, technologies like the steam engine, electricity and computers have fundamentally changed the organization of economic activity. In the future, advances in artificial intelligence and robotics hold the promise to do so once again. When such innovations are biased towards specific skills, inequality can rise fast (Katz and Murphy, 1992; Violante, 2002; Costinot and Vogel, 2010). The adjustment of labor markets, however, can take decades, as workers reallocate (Dix-Carneiro, 2014; Burstein et al., 2019; Buera et al., 2021) and younger generations acquire new skills (Chari and Hopenhayn, 1991; Porzio et al., 2020).

But do economies adjust slowly to certain technological innovations and more rapidly to others? Studies on this question are scarce. The literature has mainly focused on explaining facts that are common across episodes of adjustment, as opposed to how different technologies shape them.¹ This is surprising for two reasons. First, much evidence shows that technologies differ remarkably both in their impact on workers (Goldin and Katz, 1998, 2009; Acemoglu and Autor, 2011) and speed of diffusion (Mansfield, 1961; Rosenberg, 1972; Jovanovic and Rousseau, 2005; Comin and Mestieri, 2018). Second, understanding which innovations will trigger slower transitions is central to weighing their benefits — which may accrue mainly to future generations — against their distributional consequences, and for governments to properly design policies (Guerreiro et al., 2017).

In this paper, we show that technological transitions are not all alike. Labor markets adjust slowly to some technological innovations and more rapidly to others. We argue that one important reason is that certain innovations may benefit production activities requiring skills which are rather different from those used by incumbent workers in the rest of the economy. When such *skill specificity* is stronger, incumbent workers find it harder to reallocate, and the relative wage of activities that benefited from the innovation increases more. As a result, younger generations of workers entering the labor market have stronger incentives to enter these activities.² Therefore, a technological transition is slower when skill specificity is stronger because it is driven less by the fast reallocation of older incumbent workers and more by the gradual entry of younger generations.

To support this argument, we first document that the U.S. labor market adjusted differently to the arrival of Information & Communications Technologies (ICT) in the late twentieth century than it did to innovations in manufacturing at the beginning of that century. We then build an overlapping generations model of technological transitions. It allows us to sharply characterize the effects of skill specificity on equilibrium dynamics,

¹See Geroski (2000) for a survey of models on the S-curve of diffusion (Griliches, 1957), Helpman (1998) for a review on the diffusion of general purpose technologies, and Herrendorf et al. (2014) for facts about structural transformation.

²Younger generations can do so, for example, because they can acquire skills more easily or face lower mobility costs compared to older workers (Lagakos et al., 2018; Kambourov and Manovskii, 2008; Humlum, 2019).

match the evidence in a parsimonious way, and study its welfare implications.

We begin in Section 2 by showing that the adjustment to ICT innovations was slower and more unequal in comparison to the adjustment to the manufacturing innovations of the early 1900s. Specifically, we estimate the dynamic responses of the relative employment and wages of occupations that benefited more from the main new technologies in each episode; namely, those with a higher initial share of employment in ICT-intensive industries or manufacturing industries. We find that the ICT transition saw a fast increase in the relative wage of impacted occupations, which was followed decades later by a slow increase in their relative employment.³ In contrast, relative employment in manufacturing-intensive occupations increased fast in the first decades of the twentieth century, whereas relative wages did not change much.⁴ The joint dynamics of employment and wages suggest an increase in labor demand in ICT-intensive occupations that was slowly matched by a gradual increase in labor supply later on, while labor supply quickly responded to labor demand increases in manufacturing-intensive occupations.

Then, we show that the limited and delayed adjustment in labor supply during the ICT transition is explained by the muted reallocation of older incumbent generations. In particular, we estimate employment responses for worker groups representing younger and older generations. We find that both older and younger workers reallocated to manufacturing-intensive occupations to a similar extent shortly after 1900. However, older workers did not move towards ICT-intensive occupations. The slow expansion of these occupations after 2000 was entirely driven by the gradual entry of younger generations.⁵ These patterns are perhaps surprising given that many sources of mobility frictions likely were more, not less, severe in the early twentieth century.⁶

A possible explanation is that skill specificity was stronger in the ICT transition: older incumbent workers in other occupations had skills which were more transferable to manufacturing-intensive occupations than they were to ICT-intensive ones. Indeed, our last piece of evidence shows that stronger skill specificity helps to explain the slower and more unequal adjustment to ICT innovations. We document that ICT-intensive occupations required workers to perform tasks that were markedly different from those in the rest of the economy, whereas manufacturing-intensive ones required relatively more

³This is consistent with the evidence that skill demand rose fast in the 1980s but it then slowed down (Acemoglu and Autor, 2011), and that the ICT impact on employment polarization intensified after 2000 (Autor and Dorn, 2013).

⁴This is consistent with evidence that inequality increased more after 1980 than it did in the early 1900s (Goldin and Katz, 2009; Piketty and Saez, 2003).

⁵We also show that there were no relative wage differences across generations, suggesting no differential demand for younger workers conditional on their occupation. Moreover, for several countries, we show that employment growth in ICT-intensive occupations was stronger for younger workers than older workers. Finally, we use quasi-random regional variation in the rollout of DSL in Germany, a specific form of ICT, to show that, in regions receiving DSL early, employment in ICT-intensive occupations increased for younger workers but not for older ones.

⁶For instance, search and matching frictions were more important (Martellini and Menzio, 2020), moving to manufacturing jobs in the early 1900s often involved a costly move from rural to urban areas (Michaels et al., 2012), and better social and private insurance improved consumption smoothing during job transitions (Chetty, 2008).

similar tasks. We then estimate how the responses of relative employment and wages in more exposed occupations vary with an occupation’s task distance from the rest of the economy. We find that a higher task distance mutes the relative employment increase in more exposed occupations but amplifies their relative wage increase.⁷

Motivated by the evidence, Section 3 presents a model of technological transitions. Firms use two technologies to produce a final good. There is a continuum of worker skill types. A type determines a worker’s effective labor units when employed in each technology. This allows us to formalize the notion of skill specificity as a type’s difference in productivity between the two technologies. There are overlapping generations of workers, and each generation forms a large household. The household assigns workers to technologies by comparing their wage in each one, similar to Roy (1951). This allows the model to reproduce the reallocation of older incumbent workers following changes in relative wages across technologies. Lastly, at birth, a household chooses how much labor of each skill type to supply, given the future path of relative wages. This allows the model to replicate differences across generations in relative employment when generations face distinct future paths of relative wages at birth.

In Section 4, we turn to the dynamic adjustment of our economy to an innovation that permanently increases the productivity of one of the technologies, therefore raising its labor demand. We characterize the equilibrium adjustment dynamics by first establishing an equivalence result. To a first order approximation, the equilibrium dynamics behave *as if* they were generated by a reduced-form model of relative labor demand and supply, where the relative labor supply of each generation combines time- and generation-specific components. The generation-specific term captures the effect of changes in the present discounted value of the relative wage (which we dub q) that a generation faces at birth. As such, it governs the relative labor supply elasticity at longer horizons. In the structural model, this elasticity is higher when the cost of supplying skills is lower. The time-specific term captures the impact that changes in the relative wage at a point in time have on the relative labor supply of every generation. As such, it governs the short-run elasticity of relative labor supply. In the structural model, this short-run elasticity is lower when skill types are more different in terms of their productivity in each technology (i.e., when skill specificity is stronger).

The adjustment dynamics follow the economic logic of the seminal q -theory (Hayashi, 1982), albeit applied to labor markets. On impact, the technological innovation raises the relative wage, which induces the reallocation of older incumbent generations towards the improved technology. The increase in current and future relative wages (summarized in q) causes entering generations of households to increase their relative labor

⁷These findings add to a recent literature showing that gross job-to-job flows are smaller between occupations that require more different tasks (Gathmann and Schönberg, 2010; Traiberman, 2019; Macaluso, 2022; Schubert et al., 2021).

supply compared to older incumbents. Through the lens of the structural model, this gap results from younger generations choosing to supply more labor of skill types that are more complementary to the improved technology. Along the transition, as younger generations replace older ones, the economy's relative employment slowly increases and the relative wage declines. One interpretation of this labor reallocation process is that innovations slowly diffuse as more workers employ them over time.

Section 5 establishes our main theoretical result by comparing the adjustment in two economies with varying short-run labor supply elasticities. We interpret this as a comparison between two episodes: one in which the innovation improves a technology intensive in skills which are markedly different from those used by the rest of the economy (skill specificity is strong) and another one in which they are more similar (skill specificity is weak). Stronger skill specificity leads to a slower adjustment process, in which the increase in relative employment is more back-loaded and the rise in the relative wage is more front-loaded. On impact, stronger skill specificity *directly* weakens the reallocation of older incumbent workers. This results in a stronger increase in the present discounted value of the relative wage (q), which in turn *endogenously* strengthens the incentives of younger generations to increase their relative labor supply. Ultimately, transitional dynamics in relative employment and wages become more important because the adjustment is driven more by the slow entry of young generations.

Returning to our evidence, Section 6 shows that our parsimonious model can *quantitatively* match the different adjustment dynamics following both ICT and manufacturing innovations purely as a consequence of varying the short-run elasticity of relative labor supply (common to every generation). In particular, our baseline calibration sets this elasticity to zero, in line with the lack of reallocation of older workers during the ICT transition, and feeds a labor demand innovation of a magnitude that matches the cumulative relative employment increase in ICT-intensive occupations after 40 years. We externally calibrate the few remaining parameters using standard values from the literature (e.g., the discount and death rates). The model successfully replicates the ICT transition's fast increase in the relative wage, delayed relative employment increase, muted reallocation of older workers, and large differences in relative employment between younger and older workers. Lastly, in line with the smaller task distance between manufacturing-intensive occupations and the rest of the economy, we repeat the analysis for a larger short-run labor supply elasticity, while keeping all other parameters fixed. The model now can reproduce the transition following manufacturing innovations; its muted relative wage change, fast increase in relative employment and reallocation of older workers, and lack of differences in relative employment between younger and older workers.

We conclude the paper with a normative analysis in Section 7. Since the model can

match the evidence on *relative* responses, this gives us some confidence that we can use it to study how skill specificity affects *aggregate* dynamics following technological innovations. We show that the welfare gains from technological innovations are smaller when the adjustment is slower due to stronger skill specificity, as aggregate consumption increases take longer to materialize and mostly benefit future generations. In terms of magnitudes, moving from a zero short-run labor supply elasticity (as in our baseline calibration for the ICT transition) to, for example, a larger elasticity of 3 implies that the welfare loss from slow transitional dynamics falls by $1/3$.

Our paper is related to a literature studying slow adjustment dynamics. Close to our paper, the within and between generations margins of adjustment have also been shown to be important for changes in U.S. employment composition (Murphy and Topel, 1987; Autor and Dorn, 2009) and wage inequality (Violante, 2002), as well as structural transformation (Hobijn et al., 2019; Porzio et al., 2020). As in our model, Violante (2002) also emphasizes skill specificity in determining the reallocation of older incumbent workers, and Porzio et al. (2020) highlights skill acquisition in explaining differences between young and old generations. We add to this literature by analyzing the dynamics of adjustment after technological innovations, as well as how skill specificity shapes them by varying the strength of the within and between generations margins.

More generally, a strand of the literature has emphasized slow changes in the supply of labor to particular sectors or occupations (Matsuyama, 1992; Heckman et al., 1998; Lee and Wolpin, 2006; Dvorkin and Monge-Naranjo, 2019; Traiberman, 2019). To the best of our knowledge, our theoretical results linking how slow the adjustment is to skill specificity is new to this literature, as is our evidence for the two technological transitions. That said, we note that stronger skill specificity would also lead to a slower adjustment in quantitative dynamic Roy models where wages are determined in equilibrium, but not in models of labor reallocation where wages are exogenous. The reason is that, in the latter models, stronger skill specificity would only dampen the reallocation of older incumbent workers, but leave differences across generations unchanged because the incentives of entering generations to change their supply of skills would remain the same.

A complementary literature has analyzed slow adjustment dynamics coming from gradual changes in labor productivity due to firm learning-by-doing (Atkeson and Kehoe, 2007), knowledge diffusion (Perla and Tonetti, 2014; Lucas and Moll, 2014), the creation of new worker tasks (Acemoglu and Restrepo, 2018), or capital accumulation (Dix-Carneiro and Kovak, 2017; Buera et al., 2020). We show that gradual changes in labor productivity along the transition cannot alone account for the differences across technological transitions that we document. In particular, they cannot jointly explain the fast rise in relative wage and the slow increase in relative employment observed during the ICT transition, neither can they speak to relative employment changes which are

larger for younger workers and weaker for occupations requiring more specific skills. In contrast, our paper highlights slow adjustment dynamics due to changes in the relative supply of labor within and between generations, and show that our parsimonious model can match the features of the two transitions that we study.

Lastly, a literature has focused on the consequences of ICT and manufacturing innovations in the U.S. [Jovanovic and Rousseau \(2005\)](#) shows that, despite similar rates of price decline for ICT innovations and Electricity (used in manufacturing), diffusion was slower and productivity gains weaker during the ICT transition. Adding to these facts, our findings show that the speed of labor reallocation was different in the two episodes too; and that varying skill specificity helps to explain why. A number of papers have shown evidence of capital-skill complementarity during both the expansion of manufacturing in the early 1900s and ICT-intensive occupations later in the twentieth century ([Goldin and Katz, 1998](#); [Autor et al., 1998, 2003](#); [Goldin and Katz, 2009](#)). Our paper also rests on the idea that these new technologies complemented certain skills, thereby raising their demand when they arrived. However, we show that manufacturing and ICT innovations impacted economic activities that differed in their task distance from other activities in the economy; varying the degree to which older incumbent workers could supply the specific skills whose demand increased and, as a consequence, how slow the adjustment was.⁸

2 Evidence from Two Technological Transitions

In this section, we document three novel facts about how exposure to technological innovations affected employment and wages across occupations in the U.S. during two periods: the introduction of manufacturing-enhancing technologies in the early twentieth century, and the arrival of ICT innovations in the latter part of that century.

2.1 Exposure to new technologies and occupational outcomes

We begin by analyzing the dynamic responses of employment and wages in occupations more exposed to the new technologies in each of the two episodes. Our focus on occupations follows a recent literature studying the labor market consequences of new technologies — for a review, see [Acemoglu and Autor \(2011\)](#). Specifically, we estimate the following specification for different years t during the period starting at t_0 in which

⁸That said, our findings do not preclude that the degree of capital-skill complementarity could also have differed across episodes. For instance, because of a lower elasticity of substitution between high and low skilled workers during the ICT transition ([Goldin and Katz, 2009](#)). Through the lens of our model, though, we show that differences in the elasticity of substitution cannot alone account for the different patterns of adjustment in relative employment and wages during both episodes.

new technologies became available:

$$\log Y_{o,t} - \log Y_{o,t_0} = \alpha_t \text{Exposure}_o + \delta_t + \epsilon_{o,t}, \quad (1)$$

where $Y_{o,t}$ is either the employment or the average wage in occupation o at t , and Exposure_o is a standardized measure of occupation o 's exposure to the main new technologies introduced during the period. Thus, α_t is the estimate of how much higher the t_0 to t growth in employment or wages was for an occupation with a one standard-deviation higher shock exposure. We weight occupations by their t_0 employment to obtain representative economy-wide estimates.

To define occupation o 's exposure in each episode, we rely on the idea that the relative demand for occupation o should increase when a higher share of its workers are initially employed in industries that intensively use the main technologies that became available:

$$\text{Exposure}_o \equiv \sum_k \ell_{o,t_0}^k S^k,$$

where ℓ_{o,t_0}^k is the share of industry k in the total employment of occupation o at t_0 , and S^k is the importance for industry k of the main new technologies of each episode.

For the early twentieth century, we define S^k as the dummy variable that equals one if k is a manufacturing industry. This captures the fact that manufacturing productivity increased substantially in this period due to several innovations such as the introduction of electricity and assembly lines (David, 1990; Gordon, 2000; Jovanovic and Rousseau, 2005). It yields the intuitive implication that the most exposed occupations were associated with blue collar manufacturing jobs, like millwrights, machinists, tool makers, and electrotypers. To address concerns that estimates capture other simultaneous shocks to manufacturing demand — for example, due to the Great Depression or the WWI effort — we attest the robustness of our conclusions to defining exposure based on the electricity intensity of manufacturing industries.

For the late twentieth century, S^k is a measure of i 's share of cost spent on products and services associated with ICT in the 1997 detailed input-output table. This measure builds on the idea that the decline in the ICT cost of the late twentieth century benefited industries using ICT intensively, and implies that the most exposed occupations include instructors, scientists, and managers. Our exposure measure is highly correlated with other measures used in the literature reviewed by Acemoglu and Autor (2011) to study the labor market consequences of recent skill-biased technological innovations — for example, measures based on the education or task requirements of occupations — and, for this reason, conclusions are similar when we use these alternative measures.

We measure occupation outcomes using a sample extracted from the IPUMS U.S. Census database of males aged 16-64 years participating in the labor force. We use

the 3-digit 1950 occupation classification for the early twentieth century, and the 3-digit occupation classification in [Autor and Dorn \(2013\)](#) for the more recent episode. Our wage measure is the average annual wage income of full-time employed individuals in each occupation. Since the U.S. Census did not collect income information before 1940, we only have comprehensive wage data for the second episode. To shed light on wage responses during the first period, we use wage indices obtained from the 1975 Historical Statistics of the United States published by the U.S. Census.⁹ Appendix [A.1](#) presents more details about the data construction.

[Figure 1](#) reports our estimates of α_t for different horizons. We set t_0 as the earliest year of the U.S. Census following the arrival of the main technologies in each episode, as defined in [Jovanovic and Rousseau \(2005\)](#) and [Comin and Mestieri \(2018\)](#): 1900 for the early episode and 1980 for the latter episode. We consider changes over each decade of the 40-year period following t_0 .¹⁰ We also report estimates for the last Census before t_0 to check for pre-trends.

Consider first the two left panels depicting our estimates for the impact of ICT innovations across occupations. The top left panel shows that higher ICT exposure was not associated with stronger relative employment until the 1990s. However, after 2000, employment slowly reallocated towards more exposed occupations. The bottom left panel shows that the path of the relative wage of ICT-intensive occupations has an inverted-U shape. Our estimates indicate that the relative wage rapidly increased between 1980 and 1990. Note that, due to the lack of employment responses before 1990, this increase is unlikely to be biased by selection based on unobserved worker attributes. After 2000, the impact on relative wages slowly declines (albeit being more imprecisely estimated). Together, the relative wage and employment responses indicate that, upon the ICT arrival in the 1980s, there was an increase in the demand for ICT-intensive occupations and thus their relative wage. But this initial relative wage spike only resulted in employment reallocation two decades later.¹¹

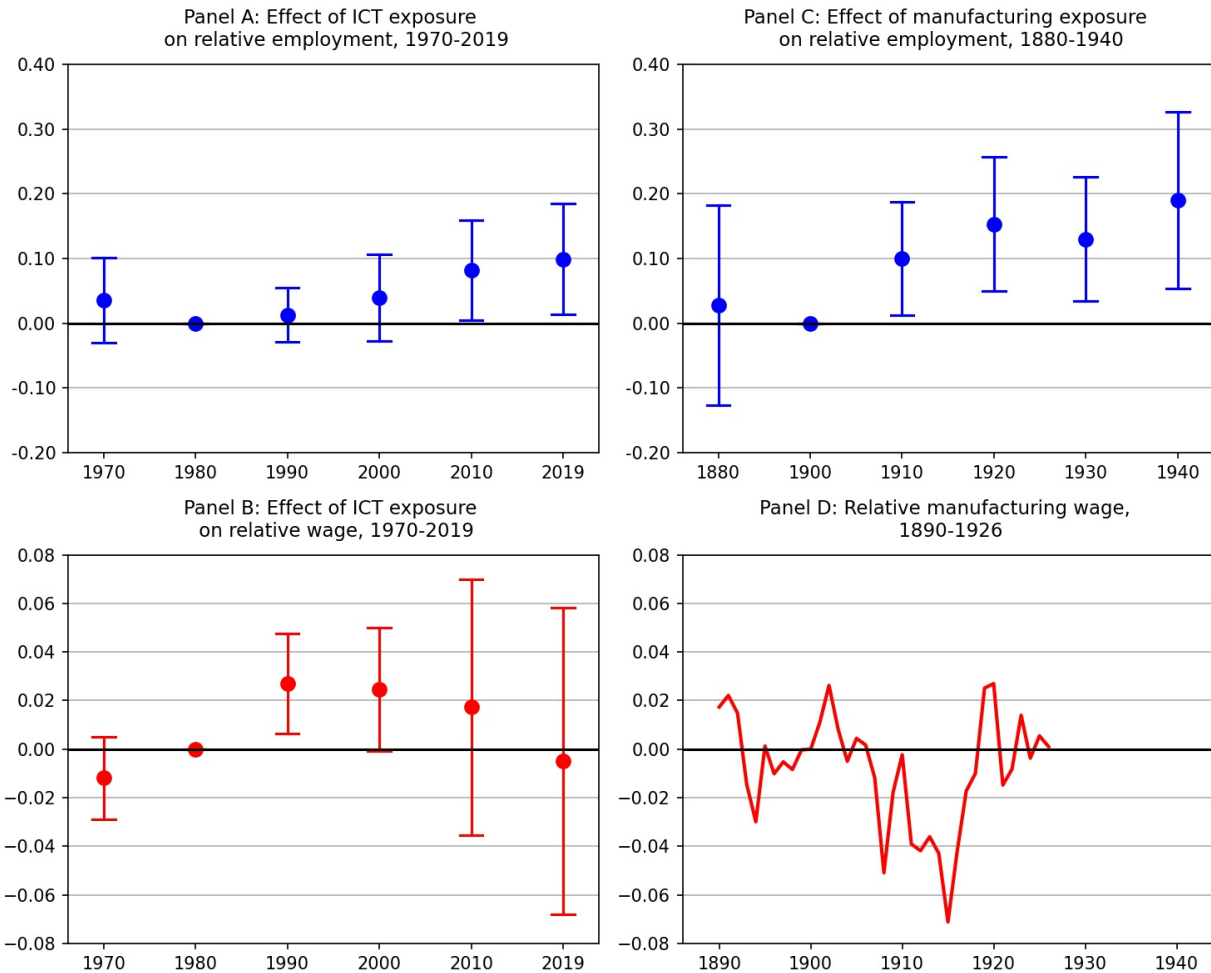
Turning to the adjustment to the innovations in manufacturing of the early 1900s, the top right panel shows that relative employment growth in more exposed occupations was both faster and stronger than that following ICT innovations in the late twentieth century. Our estimates indicate that the majority of the relative employment increase in manufacturing-intensive occupations happened during the first two decades of the

⁹This is the same primary data source in [Goldin and Margo \(1992\)](#). We define the manufacturing wage premium as the log of the ratio of wage indices for manufacturing workers (series D-781) and non-farm workers (series D-780).

¹⁰The final year of the second episode is 2019 to avoid the COVID pandemic in 2020. As in [Autor and Dorn \(2013\)](#), to increase sample size and measurement precision, we compute outcomes in 2019 by pooling the years of 2017-2019 of the American Community Survey.

¹¹Our findings are consistent with evidence reviewed in [Acemoglu and Autor \(2011\)](#). They document that the demand for high-skill workers increased fast between 1980 and the mid-1990s, but it then slows down as reflected in the weaker growth of the skill wage premium after the mid-1990s. [Autor and Dorn \(2013\)](#) show that ICT exposure induced employment reallocation across occupations in 1980-2005, with the bulk of the effect happening after 2000.

Figure 1: Exposure to Technological Innovations and Occupation Outcomes in the Two Episodes



Note. Dots are the point estimates of α_t obtained from (1) in a sample of 310 occupations for Panels A-B and 201 occupations for Panel C with positive employment in the initial and final years of the period for the U.S. Census sample of males aged 16-64yrs in the labor force. Estimates are weighted by the occupation's employment at t_0 . Vertical bars denote 95% confidence intervals clustered by occupation. The occupation exposure in Panels A-B is the occupation's average ICT-intensity across industries and in Panel C is the occupation's average manufacturing-intensity across industries. Exposure measures are normalized to have a standard deviation of one. The relative manufacturing wage in Panel D is the log of the ratio of the average wage of manufacturing workers and that of all non-farm workers in the 1975 Historical Statistics of the United States (series D-780 and D-781) — normalized to zero in 1900.

period. Over the entire period, for occupations with one standard deviation higher exposure, the relative employment growth was 0.20% in the first episode, but only 0.10% in the second. Moreover, the bottom right panel shows that the relative wage in manufacturing did not increase much during the first episode. After an early increase shortly after 1900, the manufacturing wage premium falls until WWI, and then returns to its 1900 level by the mid 1920s.¹²

In all, the joint dynamics of employment and wages suggest the following. Re-

¹²This is consistent with the evidence in Goldin and Margo (1992) that, in the decades prior to 1940, the strong relative expansion of skilled occupations happened while their wage premium was either relatively constant or moderately increasing — at least in comparison to the strong increase in the returns to skill observed in the 1980s. Similarly, Piketty and Saez (2003) find that “top income and wages shares display a U-shaped pattern over the century.”

sponses observed in the early 1900s are consistent with a labor demand increase in manufacturing-intensive occupations that was matched by a rapid increase in labor supply. In contrast, during the more recent period, wage responses suggest a rapid increase in the demand for ICT-intensive occupations in 1980-1990, which only induced labor supply responses after 2000. As such, the slower employment adjustment to the arrival of ICT in the latter episode is not consistent with stories that emphasize a slower labor demand increase — for example, because ICT or other complementary technologies diffused more slowly. The reason is that a slower labor demand increase would have implied slower increases in *both* relative wages and employment, which is not what occurred since relative wages increased decades before employment did.

Robustness. Figure A1 in Appendix A.2 investigates the robustness of these conclusions. Specifically, we consider alternative measures of occupation exposure based on electricity-intensity in the first episode, or task and education requirements in the second episode. To address concerns about confounding supply shocks across occupations, we also report estimates while using alternative sample definitions (e.g., based only on natives or white males), and controlling for job attributes (e.g., hours, self-employment). In all alternative specifications, we find similar patterns of relative employment responses for both episodes. During the ICT episode, all specifications also imply an increase in the relative wage of more exposed occupations between 1980 and 2000 while relative employment responses were absent, attesting the robustness of our conclusion of a slow supply adjustment to ICT-induced labor demand shocks. However, in line with the less precise estimates for the period after 2000, point estimates across specifications vary in the timing and magnitude of the slow down in relative wage growth.

2.2 Responses in occupation outcomes across worker generations

We established above that relative worker supply across occupations adjusted at a slower pace to the demand shock induced by post-1980 ICT innovations than to the shock induced by the innovations in manufacturing of the early 1900s. This is perhaps surprising given that many sources of reallocation frictions were likely more severe during the earlier episode. For instance, search and matching frictions declined over time (Martellini and Menzio, 2020), reallocation towards manufacturing jobs in the early 1900s also involved a costly move from traditional rural areas to new urban areas (Michaels et al., 2012), and consumption smoothing during job transitions likely improved over time due to better access to (social and private) insurance instruments (Chetty, 2008).

To investigate further what underlies the difference across the two episodes in the observed patterns of labor supply adjustment over time, we next separately estimate responses in the occupation outcomes of worker groups representing younger and older

generations. Intuitively, we would expect the overall labor supply response to be faster when generations already present in the labor market move to more exposed occupations compared to when the response is mostly driven by younger entrant generations. Specifically, we estimate the following specification:

$$\log Y_{og,t} - \log Y_{og,t_0} = \alpha_{g,t} \text{Exposure}_o + \delta_{g,t} + \epsilon_{og,t}. \quad (2)$$

The dependent variable is the log-change in either employment or wages for individuals of group g in occupation o between t_0 and t . The group-period fixed-effect $\delta_{g,t}$ absorbs any group-specific shock with a common effect on all occupations. Accordingly, $\alpha_{g,t}$ measures how much an increase of one standard-deviation in shock exposure affects occupation outcomes among workers of group g . We estimate this specification in a pooled sample of all groups, and compute standard errors clustered by occupation to account for correlated shocks affecting employment of all groups in the same occupation.

We report our baseline estimates in [Table 1](#). Column (1) of each panel replicates the respective estimate in [Figure 1](#) for the entire period in the group of all male workers aged 16-64. Columns (2) and (3) report the analog of these estimates for two worker generations: an older group, which in our baseline includes males aged 30-64, and a younger group, which includes males aged 16-29. Column (4) reports the between-generation gap implied by the pooled specification with two groups.

Panel A shows that the relative employment increase in occupations more exposed to ICT observed after 2000 was almost entirely caused by younger entrant generations. The employment impact reported in column (2) for older workers was small and non-significant. In contrast, the effect for younger workers in column (3) was 124% larger than that for older workers (while having a similar standard error). Column (4) indicates that the between-generation gap has a p-value of 0.006 and almost the same magnitude as the positive impact on the relative employment of all workers. Thus, we do not find evidence that older generations reallocated to ICT-intensive occupations in 1980-2019; instead, younger entrant generations drove the post-2000 labor supply adjustment.

One possible explanation for the stronger impact on the relative employment of younger workers is that ICT innovations only increased the demand for younger workers in more exposed occupations. However, this is not consistent with two pieces of evidence. First, Panel B indicates that higher exposure had a similar impact on relative wages for the two groups. In fact, column (4) reports a precisely estimated zero for the between-generation difference in the relative wage growth of more exposed occupations. Second, [Figure A5](#) in [Appendix A.5](#) uses on-the-job time-use data from Germany to show that, conditional on the occupation's exposure to ICT, older and younger workers report similar usage of two important ICT innovations, computers and internet.

We then turn to the earlier episode in Panel C. Estimates indicate that the strong

Table 1: Exposure to Technological Innovations and Occupation Outcomes for Younger and Older Workers in the Two Episodes

Worker group:	All 16-64yrs (1)	Older 30-64yrs (2)	Younger 16-29yrs (3)	Younger - Older (4)
<i>Panel A: Effect of ICT exposure on relative employment, 1980-2019</i>				
Exposure _o	0.099** (0.044)	0.063 (0.043)	0.141*** (0.050)	0.078*** (0.029)
<i>Panel B: Effect of ICT exposure on relative wage, 1980-2019</i>				
Exposure _o	-0.005 (0.032)	-0.003 (0.035)	0.006 (0.028)	0.009 (0.018)
<i>Panel C: Effect of manufacturing exposure on relative employment, 1900-1940</i>				
Exposure _o	0.190*** (0.069)	0.169** (0.069)	0.205*** (0.076)	0.036 (0.058)

Note. Estimates from (2) in a sample of 310 occupations for Panels A-B and 201 occupations for Panel C with positive employment in the initial and final years of the period for the U.S. Census sample of males aged 16-64yrs in the labor force. Estimates are weighted by the occupation's employment at t_0 . Dependent variable is the change in the outcome and period indicated in each panel. Estimates in column (1) based on a single group of all males aged 16-64yrs. Estimates in columns (2) and (3) based on a pooled regression (including group-period fixed effects) with two groups: Younger group includes workers aged 16-29yrs and Older group includes workers aged 30-64 yrs. Column (4) reports the difference between the coefficients for younger workers in column (3) and older workers in column (2). The occupation exposure in Panels A-B is the occupation's average ICT-intensity across industries and in Panel C is the occupation's average manufacturing-intensity across industries. Exposure measures are normalized to have zero mean and unit standard deviation. Robust standard errors in parentheses are clustered by occupation with *** denoting a p-value <0.01 and ** a p-value <0.05.

employment growth in 1900-1940 of manufacturing-intensive occupations was driven by similar patterns of reallocation for older and younger workers. Columns (2) and (3) report large positive effects on relative employment for both worker generations. The impact for younger workers is only 21% higher than that for older workers, and their difference in column (4) is not statistically significant at usual levels (i.e., the p-value is 0.54). As [Table A1](#) in [Appendix A.3](#) shows, both younger and older generations experienced similar increases in their relative employment in manufacturing-intensive occupations between 1900 and 1920, implying that all generations contributed to the rapid employment reallocation documented in [Panel C](#) of [Figure 1](#). Thus, all worker generations reallocated simultaneously towards manufacturing-intensive occupations shortly after their relative demand increased in the early twentieth century.¹³

Our results suggest that the adjustment in labor supply across occupations in response to the demand shock triggered by ICT innovations was limited and delayed due to the muted reallocation of older generations. In fact, the post-2000 labor supply adjustment in ICT-intensive occupations was almost entirely driven by the response in the relative employment of younger entrant generations. However, in the early twentieth

¹³We note that the lack of wage microdata before 1940 prevents us from reporting the manufacturing wage premium separately for older and younger workers — that is, the group-specific analog of [Panel D](#) in [Figure 1](#).

century, both younger and older generations contributed to the fast labor supply adjustment to the shock in the relative demand for manufacturing-intensive occupations. These patterns are consistent with two explanations. The first is that mobility frictions were less severe in the early 1900s (especially for older workers), regardless of which occupations workers would move towards. However, as mentioned above, many sources of mobility frictions likely were more, not less, severe in the earlier episode. So, while we cannot entirely discard it, we do not find this first explanation appealing. The alternative explanation is that it was easier for older incumbent workers to reallocate towards manufacturing occupations than it was for them to reallocate towards ICT-intensive occupations. This could be, for instance, because their skills were more transferable to the manufacturing-intensive occupations that were expanding in the early episode. The next section provides evidence that this mechanism helps to explain the differences across the two transitions.

Robustness. Appendix A.3 investigates the robustness of the results in this section. Note first that the estimates of the between-group gap in column (4) obtained from our fully saturated pooled specification are identical when we introduce occupation-time fixed-effects that capture the common impact of any occupation-specific shock on all generations. Second, our results are qualitatively robust to the age cutoff used to define the two worker generations. However, the stronger reallocation of younger workers to ICT-intensive occupations becomes more pronounced as we decrease the age cutoff defining the young group (that is, as we decrease the average age of the young group). This is not true for relative employment in manufacturing-intensive occupations in the early 1900s. Finally, our results are similar when we use the same alternative exposure measures and sample definitions introduced in the previous section.

Decomposition of employment reallocation across cohorts. An alternative to estimating responses by age groups is to implement a cohort decomposition of the transition. While somewhat harder to interpret, we do find that our conclusions are robust to such an approach. In Table A2 of Appendix A.3, we extract the common and cohort-specific components of the reallocation towards highly exposed occupations over the decades of each episode. Our results show that relative employment growth in highly exposed occupations was entirely driven by cohort-specific effects in the more recent ICT-driven episode, but period effects common to all cohorts explain almost entirely the reallocation towards manufacturing-intensive occupations in the early episode.

Cross-occupation evidence from other countries. In Figure A3 of Appendix A.3, we use cross-country data on employment by broad occupation and age groups to document

that, in all seventeen developed countries in our sample, relative employment growth in ICT-intensive occupations for younger workers was stronger than that for older workers.

Evidence from quasi-random variation in broadband internet arrival across German regional labor markets. In Appendix A.5, we provide additional causal evidence on the impact of ICT innovations across occupations and generations in Germany. We first use comparable micro-data data to show that employment growth was higher in occupations more exposed to ICT, and that such an effect was stronger for younger workers than for older workers. We then focus on how the German labor market responded to one specific ICT innovation, broadband internet. Our strategy leverages the fact that the rise in penetration of broadband internet in Germany was fast in the early 2000s, but there was quasi-random regional variation in broadband penetration (Falck et al., 2014). We establish that the arrival of broadband internet in a region caused younger workers to reallocate towards ICT-intensive occupations, but it did not affect the allocation of older workers across occupations.

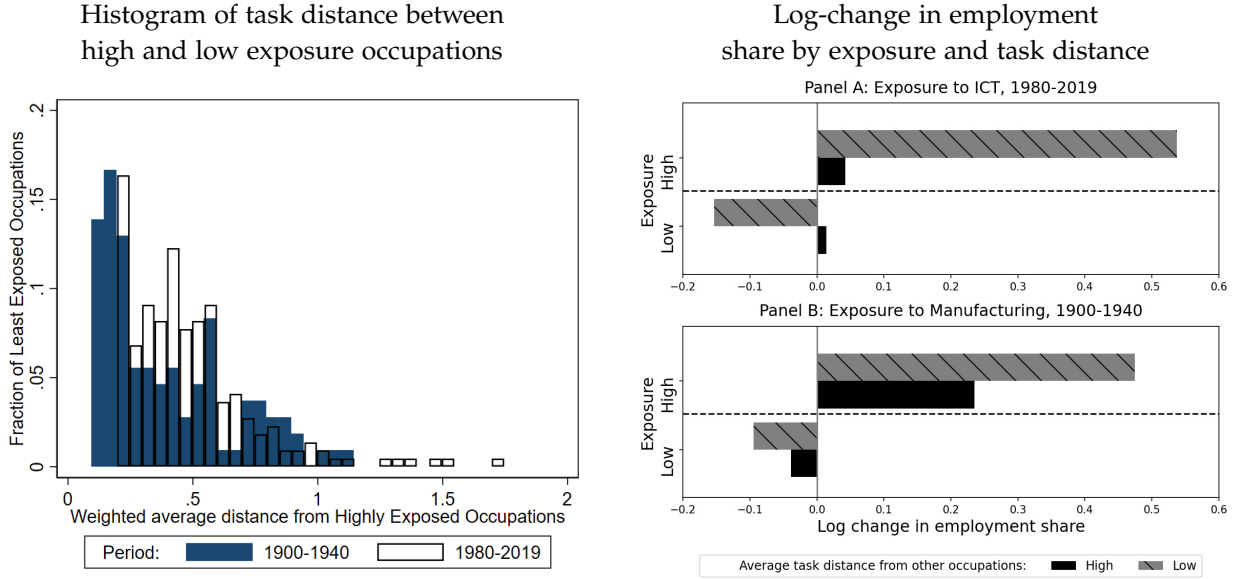
2.3 Differences in task content between occupations

We now investigate how the two episodes differed in terms of the ability of older incumbent workers in the rest of economy to supply the skills required by the occupations whose demand increased. We do so by measuring, for each episode, how different the task requirements of the occupations in high demand were in comparison to those of other occupations. Our approach builds on a recent literature showing that, both in the U.S. and other countries, gross job-to-job flows are higher between occupations that require similar tasks (Gathmann and Schönberg, 2010; Traiberman, 2019; Macaluso, 2022; Schubert et al., 2021). Intuitively, task requirements capture the common component of the activities that workers need to perform on the job and, consequently, reallocation should be easier across occupations requiring workers to perform similar activities (for a given change in relative wages). Formally, for any two occupations o and o' , we measure the distance between the distribution of their task requirements using an entropy metric: $D_{o,o'} = \sum_{s \in \mathcal{S}} \sigma_{o,s} \log \sigma_{o,s} / \sigma_{o',s}$, where $\sigma_{o,s}$ is occupation o 's intensity in task s .¹⁴ We follow Autor and Dorn (2013) to measure $\sigma_{o,s}$ for $s \in \mathcal{S} \equiv \{Manual, Routine, Cognitive\}$ among the occupations in our sample in each of the two episodes.¹⁵

¹⁴This is a standard metric of the distance between two distributions. Results are similar when using alternatives such as the Euclidean metric.

¹⁵By using the same occupation task content in the two periods, we implicitly assume that the task intensity ranking across occupations in 1900 has a high correlation with that of 1980. We do so for lack of comprehensive task content data for the early twentieth century. Our approach is reasonable because, as Spitz-Oener (2006) shows, the task intensity ranking across occupations did not change much over time (as the majority of changes happened within occupations).

Figure 2: Exposure, Task Distance, and Employment Reallocation in the Two Episodes



Note. The left panel of this figure presents the histograms of the average task distance from high-exposure occupations in \mathcal{O}_H across low-exposure occupation in \mathcal{O}_L , which are respectively defined as the top 25 percent and the bottom 75 percent of occupations in terms of exposure to manufacturing (1900-1940) or ICT intensive industries (1980-2019). The right panel shows, in each episode, the log-change in the employment share on four groups of occupations defined in terms of shock exposure (“High” exposure in \mathcal{O}_H and “Low” exposure in \mathcal{O}_L) and average task distance from other occupations (the “High” group includes occupations with above median average task distance from other occupations, and “Low” all other occupations).

We start by measuring the average task distance between the occupations with the highest and the lowest levels of shock exposure in each of the two episodes. We consider two sets of occupations in each episode: those in the top 25 percent (\mathcal{O}_H) and those in the bottom 75 percent (\mathcal{O}_L) of the empirical exposure distribution. We then compute for each occupation $o \in \mathcal{O}_L$ its average task distance from the occupations in \mathcal{O}_H , weighted by their employment at t_0 .

The left panel of [Figure 2](#) depicts two histograms, one for each period, of such distance measure across occupations. The key takeaway from the figure is that the task distance distribution for ICT exposure in the latter episode has more mass on higher distance values than the distribution for manufacturing exposure in the earlier episode. In fact, we obtain a p-value of 0.00 for the test of whether the distribution of distances in the latter episode first-order stochastically dominates that of the earlier episode. Thus, occupations with low and high exposure to the manufacturing innovations of the early 1900s were much more similar in their task contents than occupations with low and high exposure to the latter ICT innovations.

To further investigate the relationship between shock exposure and task distance in the two episodes, we estimate how an occupation’s exposure was associated with its (employment-weighted) average task distance from all other occupations. Column (1)

of [Table 2](#) shows that occupations with higher ICT exposure in 1980 required tasks that were significantly different from the tasks used in the rest of the economy. However, as column (4) of [Table 2](#) shows, no such relation existed between manufacturing exposure and task distance in 1900, since the estimated coefficient is close to zero (with a similar standard error).

We thus conclude that the new technologies of the early 1900s raised the relative demand for manufacturing-intensive occupations that required workers to perform activities using skills that were relatively similar to those used in other occupations. The opposite was true during the second episode: ICT technologies increased the relative demand for occupations requiring skills that were remarkably different from those used in the rest of the economy.

To the extent that relative labor supply is more elastic between occupations with a similar task content ([Traiberman, 2019](#)), this distinction between the two episodes helps to explain why the reallocation of older incumbent workers was more muted during the adjustment to ICT innovations. The right panel of [Figure 2](#) provides evidence that indeed this mechanism played a role. It shows that the employment expansion (contraction) of the occupations with the highest (lowest) levels of exposure in each episode happened almost entirely from the occupations with a similar task content to those in the rest of the economy — that is, those with a low average task distance whose employment growth is denoted by the gray bars.

We now complement this evidence by estimating how the responses of relative employment and wages in more exposed occupations vary with the occupation’s task distance from the rest of the economy:

$$\log Y_{o,t} - \log Y_{o,t_0} = \alpha_t \text{Exposure}_o + \beta_t \text{Exposure}_o \times \text{TaskDistance}_o + \delta_t + \epsilon_{o,t}. \quad (3)$$

where Exposure_o is the same exposure measure defined above, and TaskDistance_o is occupation o ’s average task distance from all other occupations, weighted by their employment at t_0 . We normalize occupation exposure and task distance to have means of zero and unit standard deviation in each episode. Hence, β_t measures how much higher or lower was the response to shock exposure in an occupation with one standard deviation higher task distance from the rest of the economy.

Columns (2) and (3) of [Table 2](#) indicate that, for an occupation with a higher task distance from others, the impact of higher ICT exposure on relative employment was weaker but the impact on relative wages was stronger. These effects are economically large: the impact of exposure on employment is close to zero for occupations with one standard deviation higher average task distance. For the early twentieth century, the point estimate in column (5) is even more pronounced, but also less precise.

These response patterns are consistent with a lower elasticity of relative labor sup-

Table 2: Exposure, Task Distance, and Responses in Occupation Outcomes in the Two Episodes

Dependent Variable:	ICT exposure			Manufacturing exposure	
	Avg. Task distance	1980-2019 log-change in employment	avg. wage	Avg. Task distance	1900-1940 change log employment
	(1)	(2)	(3)	(4)	(5)
Exposure _o	0.144*** (0.054)	0.119** (0.047)	-0.021 (0.031)	-0.006 (0.049)	0.187*** (0.040)
Exposure _o × TaskDistance _o		-0.074* (0.043)	0.049** (0.024)		-0.137* (0.071)

Note. Estimates from (3) in a sample of 291 occupations for columns (1)-(3) and 200 occupations for columns (4)-(5) with positive employment in the initial and final years of the period for the U.S., and task content information. Census sample of males aged 16-64yrs in the labor force. Estimates are weighted by the occupation’s employment at t_0 . Dependent variable indicated in each column. The occupation exposure in columns (1)-(3) is the occupation’s average ICT-intensity across industries and in columns (4)-(5) is the occupation’s average manufacturing-intensity across industries. The occupation “Task Distance” is its average task distance from all other occupations, weighted by those occupations employment at t_0 . Exposure and Task Distance are normalized to have zero mean and unit standard deviation. Robust standard errors in parentheses are clustered by occupation with *** denoting a p-value <0.01, ** a p-value <0.05, and * a p-value <0.10.

ply in occupations requiring tasks that are more different from those used in the rest of the economy. Together with the fact that ICT-intensive occupations had a higher task distance from the rest of economy, we show below that this can jointly explain the different employment responses of older and younger workers in the two episodes (as documented in Section 2.2), and, in turn, the different dynamic adjustments of overall relative employment and wages (as documented in Section 2.1). Finally, while our empirical findings alone do not speak to aggregate outcomes, we show that, when combined with the model below, the distinct observed patterns of adjustment imply remarkably distinct welfare gains from the new technologies introduced in the two episodes.

Robustness. In Appendix A.4, we establish that the takeaway from Figure 2 is robust to alternative choices of distance metric, exposure measures, weighting schemes, and thresholds for defining highly exposed occupations. For all alternative specifications, the distribution of task distances between the most and least exposed occupations to ICT in latter episode first-order stochastically dominates that implied by manufacturing exposure in the early 1900s.

Training in ICT-intensive occupations and relative employment of younger entrant generations. One natural question is why younger entrant generations were able to increase their relative employment in ICT-intensive occupations. In Appendix A.5, we shed light on this question by leveraging Germany’s unique large-scale vocational training program. We show first that higher ICT exposure was associated with stronger growth in the number of young trainees across occupations. Second, we use again the same source of quasi-random regional variation in broadband arrival to show that there was a spike in vocational training in more exposed occupations prior to the reallocation of younger workers into these occupations. In terms of magnitude, the impact of exposure

on the relative employment of younger generations was similar to the impact on the relative number of trainees. These findings indicate that acquisition of occupation-specific human capital by young workers contributed to the observed between-generation gap in employment responses to ICT innovations.

3 A Model of Technological Transitions

We now present a model of technological transitions that can generate the patterns of adjustment across occupations and generations documented above for the two episodes. We consider a closed economy in continuous time. There is a single final good whose production uses two intermediate inputs, the high-tech ($k = H$) and the low-tech ($k = L$) goods. Input production uses only labor. There is a continuum of worker skill types $i \in [0, 1]$. The worker's skill type determines her effective labor supply when employed with the H and L production technologies. There are overlapping generations of workers, and each generation forms a large household. The household consumes the final good, assigns workers into a technology, and chooses how much labor of each type to supply at birth. Labor and good markets are competitive.

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

$$Y_t = \left[(A_t X_{Ht})^{\frac{\theta-1}{\theta}} + (X_{Lt})^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}} \quad (4)$$

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

A skill type determines the effective labor units that a worker supplies when employed in the production of each of the two inputs.¹⁶ Specifically, the production functions of L and H are respectively

$$X_{Lt} = \int_0^1 \alpha(i) s_{Lt}(i) di, \quad (5)$$

$$X_{Ht} = \int_0^1 \alpha(i) \sigma(i) s_{Ht}(i) di, \quad (6)$$

where $s_{kt}(i)$ is the density of workers of type i employed with technology k at time t , $\alpha(i)$ is the number of efficiency labor units of type i , and $\sigma(i)$ is i 's differential number of efficiency units in high-tech production. Therefore, $\alpha(i)$ governs vertical differences in productivity across skill types (i.e., *absolute advantage*), whereas $\sigma(i)$ governs horizontal productivity differences across technologies for a given type (i.e., *comparative advantage*).

¹⁶Consistent with the evidence in the previous section, we assume that the productivity of a worker when operating any of the two technologies only depends on their skills and not on their generation.

Without loss of generality, we assume that $\sigma(i)$ is increasing in i , so that the skills of higher i types are more complementary to the H technology than to the L technology.¹⁷

Let w_{kt} denote the wage rate of one efficiency unit of labor employed with technology k at time t . In a competitive environment in which firms take wages as given, the labor earnings of type i when employed with the H and L technologies are respectively

$$w_{Ht}(i) = \omega_t \sigma(i) \alpha(i) \quad \text{and} \quad w_{Lt}(i) = \alpha(i), \quad (7)$$

where $\omega_t \equiv w_{Ht}/w_{Lt}$ is the relative wage per efficiency unit in high-tech good. In the rest of the paper, we refer to ω_t simply as the relative wage, and specify the economy's numeraire as $w_{Lt} \equiv 1$.

Under perfect competition, ω_t is also the relative price of the high-tech input faced by final producers. Thus, their cost minimization problem implies that the relative demand for effective labor units in H (or simply the relative demand in H) is

$$x_t \equiv \frac{X_{Ht}}{X_{Lt}} = \omega_t^{-\theta} A_t^{\theta-1}, \quad (8)$$

and their zero profit condition implies that the final good price is

$$P_t = (1 + \omega_t x_t)^{\frac{1}{1-\theta}}. \quad (9)$$

Workers. We consider overlapping generations of workers whose birth and death follow a Poisson process with rate δ . We assume that each generation τ forms a large household at birth.¹⁸

At each time t , the household chooses how to assign workers to technologies and uses their labor income to buy the final good. The optimal assignment maximizes static labor income, so that a worker of skill type i earns

$$w_t(i) = \max\{\omega_t \sigma(i), 1\} \alpha(i). \quad (10)$$

At birth, the household of generation τ chooses how much labor of each skill type to supply. It does so to maximize the present discounted value of total household log-consumption, net of the utility cost from labor supply. Specifically, the problem of the

¹⁷Appendix C.1 provides one economic interpretation for a skill type. We show that (5)–(6) arise when production combines individual-level output of each worker's "cognitive" and "non-cognitive" task input. Differences in productivity across types are a consequence of differences in their relative ability to perform cognitive tasks, and differences in the importance of cognitive tasks for the two technologies.

¹⁸This assumption implies that technological innovations will only have distributional consequences across generations in terms of welfare. In Adão et al. (2020), we explored an alternative model where individuals workers chose lotteries over skill types. As such, they were exposed to ex-post risk from the realization of the lottery, thus creating welfare differences within generations. This alternative model, however, delivers identical predictions for the approximate equilibrium dynamics we characterize in the following sections.

household of generation τ is

$$\begin{aligned}
U_\tau \equiv & \max_{\tilde{s}_\tau(i), C_{\tau,t}, L_\tau} \int_\tau^\infty e^{-\rho(t-\tau)} \log(C_{\tau,t}) dt - L_\tau \\
& \text{s.t.} \\
P_t C_{\tau,t} = & \int_0^1 w_t(i) \tilde{s}_\tau(i) di \\
L_\tau = & \left(\int_0^1 \tilde{s}_\tau(i)^{\frac{1}{\nu}+1} di \right)^{\frac{1}{\frac{1}{\nu}+1}} \leq 1
\end{aligned} \tag{11}$$

where $\tilde{s}_\tau(i)$ is the labor supply of type i from generation τ , $C_{\tau,t}$ is consumption at time t of generation τ , and L_τ is the total labor supply which we normalize to one.

Discussion. At this point, it is worth discussing how three features of the model relate to the evidence in [Section 2](#). First, skill heterogeneity coupled with the skill-technology assignment in (10) implies that, when faced with different relative wages over time, each generation will choose to reallocate workers across technologies (given their skills). This allows the model to rationalize the type of within-generation worker reallocation that we documented for the manufacturing innovations of the early twentieth century. Second, the skill supply decision of entering generations implies that, when faced with different future paths of relative wages upon entry, different generations will choose different skills to supply and will, therefore, allocate their workers differently across technologies.¹⁹ This allows our model to rationalize the type of between-generation differences in employment reallocation that we documented for the ICT innovations of the late twentieth century. It is also consistent with the observed increase in vocational training in ICT-intensive occupations by younger workers in Germany. Finally, as discussed in more detail below, the slope of $\sigma(i)$ measures how different skill types are in terms of relative productivity across technologies and, thus, controls the elasticity of relative labor supply of incumbent workers. This will allow us to study how different the economy's behavior is when the *distance* between skill requirements across technologies is different, such as the differences across the two episodes documented in [Section 2.3](#).

Our preferred economic interpretation of these assumptions is that changes in relative wages induce older workers to move towards sectors or occupations that require similar skills and thus entail minimal retraining. Intuitively, in contrast to younger workers, they face a high cost of fundamentally changing career paths by acquiring completely different skills. For tractability, we collapse into a one-time decision upon entry the skill investment that in reality occurs through formal schooling, major college choice, on-the-

¹⁹Our baseline model assumes that only new generations choose the supply of specific skills and that this decision is independent of the skills from previous generations. We relax both assumptions in [Section 6.4](#), allowing older generations to re-train and new generations to learn-from-others when choosing their skills supply.

job learning, or vocational training. Note, however, that this decision goes beyond traditional unidimensional decisions to acquire more human capital (or whether to attend college) that allow workers to vertically differentiate. Instead, it is a multidimensional decision to supply different types of skills, driven both by vertical considerations — expressed in absolute advantage $\alpha(i)$ — as well as horizontal considerations — expressed through comparative advantage $\sigma(i)$ and relative wages ω_t .

Equilibrium. Given our overlapping generations structure, the aggregate labor supply of a skill type, $s_t(i)$, follows the Kolmogorov-Forward equation,

$$\frac{\partial s_t(i)}{\partial t} = -\delta s_t(i) + \delta \tilde{s}_t(i). \quad (12)$$

Finally, an equilibrium must satisfy market clearing for all t . By Walras' law, it suffices that the relative demand for effective labor in H production in (8) is equal to the relative supply of effective labor units in H , x_t implied by the ratio of (5)–(6) under the equilibrium assignment:

$$x_t = \omega_t^{-\theta} A_t^{\theta-1} = \frac{\int_{G_t(i)=H} \alpha(i) \sigma(i) s_t(i) di}{\int_{G_t(i)=L} \alpha(i) s_t(i) di} \quad (13)$$

where $\{G_t(i) : i \in [0, 1] \rightarrow \{H, L\}\}$ is the equilibrium technology-skill assignment implied by (10). In the rest of the paper, we refer to the equilibrium x_t as the relative employment of effective labor units in H or, simply, as relative employment.

Definition 1 (Competitive Equilibrium) *Given an initial $s_0(i)$ and a path for $\{A_t\}_{t \geq 0}$, 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 type supply $\{s_t(i)\}_{t \geq 0}$, a generation τ 's skill type supply $\{\tilde{s}_\tau(i)\}_{\tau \geq 0}$, the relative wage, employment and final price index $\{\omega_t, x_t, P_t\}_{t \geq 0}$, such that*

1. *Given $\{\omega_t\}_{t \geq 0}$, $\{G_t(i), \tilde{s}_t(i)\}_{t \geq 0}$ are determined by (10) and (11).*
2. *Given $s_0(i)$ and $\{\tilde{s}_t(i)\}_{t \geq 0}$, $\{s_t(i)\}_{t \geq 0}$ is determined by (12).*
3. *For all $t \geq 0$, the market clearing condition (13) is satisfied and P_t is given by (9).*

In principle, the characterization of the equilibrium dynamics involves solving a complex infinite-dimensional fixed-point problem. To see this, consider a conjectured path for the relative wage $\{\omega_t\}_{t \geq 0}$. This path determines the skill supply decisions of new generations $\{\tilde{s}_\tau(i)\}_{\tau \geq 0}$ from (11) and, consequently, the path for the aggregate labor supply of skill types $\{s_t(i)\}_{t \geq 0}$ from (12) given $s_0(i)$. The relative wage path also determines the assignment of workers from (10). Taken together, the skill supply and assignment decisions determine the relative supply of effective labor units in H . In an equilibrium, the implied relative supply needs to be equal to relative demand at the conjectured path

for the relative wage — i.e., they need to be consistent with market clearing in (13). The next section shows that, up to a first order approximation, the equilibrium dynamics can be sharply characterized.

4 Equilibrium Dynamics

In this section, we begin by showing that, up to a first order, the equilibrium dynamics of relative employment and wages in our structural model are equivalent to those from a reduced-form model of the relative labor supply and demand across technologies. We then use this characterization to study the economy's adjustment to a technological innovation.

4.1 A reduced-form equivalence

In order to provide an analytical characterization of the equilibrium dynamics, we consider a log-linear expansion around the stationary equilibrium.²⁰ We let " $\hat{\cdot}$ " denote variables in log-deviations from their levels in the stationary equilibrium. The following proposition establishes our main equivalence result for equilibrium dynamics in relative employment and wages.

Proposition 1 (Reduced-form equivalence) *The equilibrium dynamics of relative employment and wages $\{\hat{x}_t, \hat{\omega}_t\}$ are identical to those from a reduced-form model where relative labor demand at time t is*

$$\hat{x}_t = -\theta\hat{\omega}_t + (\theta - 1)\hat{A}_t, \quad (14)$$

and relative labor supply at time t is a population-weighted average of the labor supply of different generations $\tau \leq t$ at time t , $\hat{x}_{\tau,t}$,

$$\hat{x}_t = \delta \int_{-\infty}^t e^{-\delta(t-\tau)} \hat{x}_{\tau,t} d\tau, \quad (15)$$

where

$$\hat{x}_{\tau,t} = \eta\hat{\omega}_t + \psi\hat{q}_\tau \quad \text{with} \quad \hat{q}_\tau \equiv \int_{\tau}^{\infty} e^{-\rho(s-\tau)} \hat{\omega}_s dt. \quad (16)$$

Proof. See [Appendix B.1](#). ■

The proposition shows that relative labor supply across technologies in the structural model of [Section 3](#) behaves *as if* it was coming from a reduced-form model in which the

²⁰In [Adão et al. \(2020\)](#), we also numerically solve for the equilibrium and show that our main insights are not driven by the equilibrium approximation.

relative labor supply of each generation combines time- and generation-specific components. These components determine how the relative supply elasticity varies over different horizons.

The generation-specific term, $\psi \hat{q}_\tau$, is a forward-looking component capturing the impact that changes in the present discounted value of the relative wage at birth (\hat{q}_τ) have on a generation's relative labor supply, permanently. In the reduced-form model, ψ , together with δ , thus govern the elasticity of relative labor supply over long horizons. In the structural model, this generation-specific component arises from the skill supply choices at birth. In particular, ψ is inversely proportional to the cost of adjusting the skill supply as governed by $1/\nu$ (i.e., $\psi \propto \nu$). Intuitively, when this cost is higher, skill supply is less responsive to changes in the path of future relative wages, and so the elasticity of labor supply at longer horizons is smaller.

The time-specific term, $\eta \hat{\omega}_t$, captures the instantaneous impact that changes in the relative wage ($\hat{\omega}_t$) have on the relative labor supply of every generation. In the reduced-form model, η thus governs the short-run elasticity of relative labor supply. In the structural model, this short-run elasticity is intrinsically linked to how similar skill types are in terms of their productivity in each technology (as governed by the function $\sigma(i)$). Formally,

$$\eta \propto \left(\frac{\partial \log \sigma(i)}{\partial \log i} \Big|_{i=l} \right)^{-1}$$

with l denoting the marginal skill type that earns the same wage in both technologies in the stationary equilibrium (as implicitly defined by (10)). We will say that *skill specificity* is stronger when $\partial \log \sigma(i) / \partial \log i |_{i=l}$ is larger because a worker's productivity changes more when deployed to the H - rather than to the L -technology. Intuitively, when skill specificity is stronger, skill types are less similar and workers become less transferable across technologies. As such, the household requires larger changes in relative wages to switch them across technologies, and the short-run elasticity of relative labor supply η is lower. In what follows, we thus associate stronger skill specificity with a lower η .

The proof of the proposition shows that, up to a first order approximation, the elasticity $\partial \log \sigma(i) / \partial \log i |_{i=l}$ summarizes the heterogeneity between marginal types around the threshold l and, thus, how skill heterogeneity shapes the equilibrium dynamics of relative labor supply x_t . The reason is that most workers never switch technologies along an equilibrium path in which the relative wage remains around its stationary level. Therefore, when evaluating changes in relative labor supply over time, the effect of changes in the supply of the marginal types that switch technologies along the transition are of second order.

4.2 The economy's adjustment to a technological innovation $\Delta \log(A)$

We are now ready to characterize the dynamic adjustment of our economy to a permanent, unanticipated increase in the relative productivity of high-tech production in the structural model of [Section 3](#) or, equivalently, an increase in the relative labor demand in the reduced-form model of [Section 4.1](#).²¹

We consider a shock $\Delta \log(A) > 0$ at $t = 0$ in an economy starting from a stationary equilibrium immediately prior to the shock at time $t = 0^-$. The next proposition characterizes the transitional dynamics of $\{\hat{x}_t, \hat{q}_t, \hat{\omega}_t\}$ after the shock. [Appendix B.2](#) completes the characterization by deriving the short-run and long-run responses.

Proposition 2 (*q-theory transitional dynamics*) *Given the initial condition \hat{x}_0 and the terminal condition $\lim_{t \rightarrow \infty} \hat{x}_t = 0$, the equilibrium transitional dynamics of $\{\hat{x}_t, \hat{q}_t, \hat{\omega}_t\}$ are described by the following system of ordinary differential equations:*

$$\frac{\partial \hat{x}_t}{\partial t} = -\delta \hat{x}_t + \frac{\theta \psi \delta}{\eta + \theta} \hat{q}_t \quad (17)$$

$$\rho \hat{q}_t = \hat{\omega}_t + \frac{\partial \hat{q}_t}{\partial t} \quad (18)$$

$$\hat{\omega}_t = -\frac{1}{\theta} \hat{x}_t. \quad (19)$$

The equilibrium $\{\hat{x}_t, \hat{q}_t, \hat{\omega}_t\}_{t \geq 0}$ is saddle-path stable with a rate of convergence λ .

Proof. See [Appendix B.1](#). ■

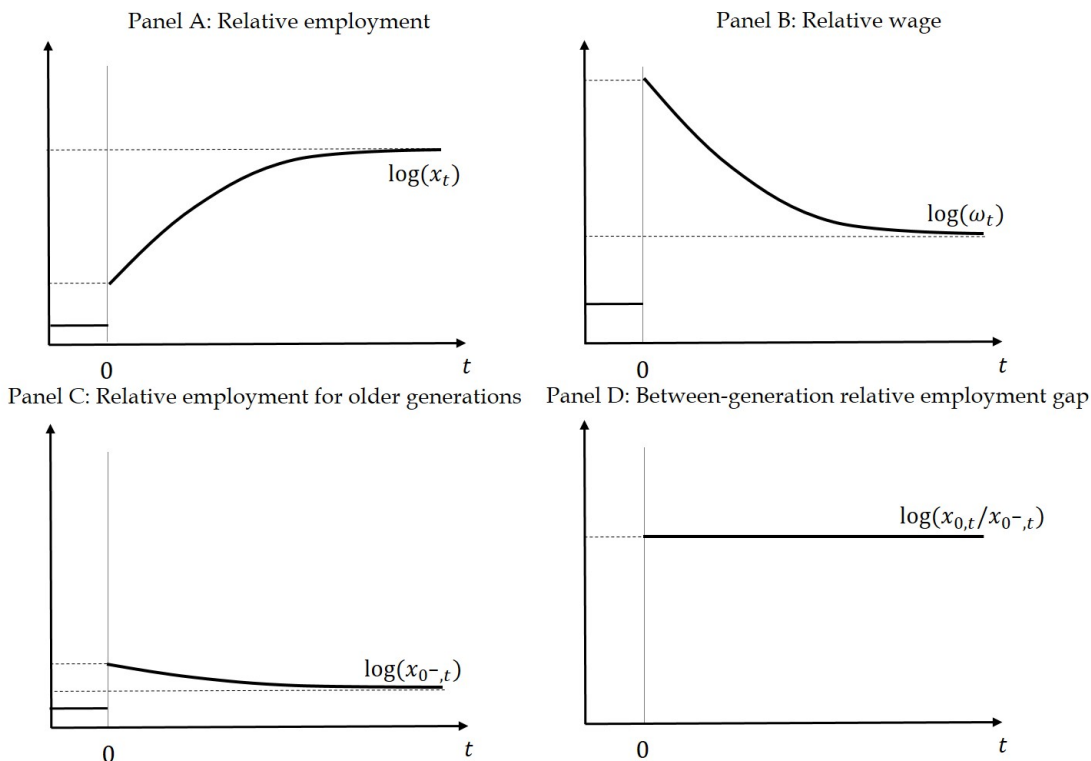
The system of ordinary differential equations is a rather standard one in macroeconomics: it has one control variable (q_t) and one predetermined variable (x_t), as well as one Kolmogorov-forward equation (17) and one Kolmogorov-backward equation (18). The system is in fact mathematically isomorphic to the one in the q -theory of capital investment. In our model, the present discounted value of relative wages \hat{q}_t represents the shadow price of the human capital “asset” associated with having one additional effective unit of labor in high-tech production. As in the q -theory, we can see in (17) that parameters governing the “costs of adjustment” in the economy (δ and ψ) affect the sensitivity of changes in the state variable ($\frac{\partial \hat{x}_t}{\partial t}$) to the control variable (\hat{q}_t). Moreover, since η and θ affect how changes in relative wages translate into change in labor supply and demand, they also mediate the impact of q_t on x_t . The last part of the proposition shows that (locally) the equilibrium exists and is unique — a consequence of saddle-path stability. Given an initial condition \hat{x}_0 , \hat{x}_t and \hat{q}_t converge at a constant rate to the

²¹We emphasize that [Section 3](#) has given a particular micro-foundation for the reduced-form model. But any other structural model that would deliver the same reduced-form model would have the same positive implications of technological innovations that we develop in [Sections 4.2 to 6](#). Having said so, the particular structural model of [Section 3](#) does matter for analyzing normative implications, as we do in [Section 7](#).

stationary equilibrium, where $-\lambda$ is the negative eigenvalue of the system of differential equations.

Figure 3 illustrates the economy's dynamic response relative to the initial stationary equilibrium after the technological shock at $t = 0$. For example, the response of relative employment is $\Delta \log(x_t) \equiv \log(x_t/x_{0^-})$. We do so for the case of a demand elasticity θ larger than 1.

Figure 3: Dynamic Responses to a Technological Innovation ($\theta > 1$)



In the short-run, the increase in the relative labor demand causes the relative wage to rise ($\Delta \log(\omega_0) > 0$). The higher relative wage induces the reallocation of existing worker generations from the L to the H technology ($\Delta \log(x_{0^-,0}) > 0$), as governed by the elasticity η . In turn, relative employment increases too ($\Delta \log(x_0) > 0$).

The persistent increase in the relative wage implies that younger worker generations entering after the shock face a higher present discounted value of the relative wage q_t compared to that faced by older generations born before the shock. This induces younger generations to increase their relative labor supply in comparison to older generations. To illustrate this, Panel D in the figure shows a positive gap $\Delta \log(x_{0,0}) > \Delta \log(x_{0^-,t})$ in relative employment between the initial generation born right after shock ($\tau = 0$) and those born before ($\tau = 0^-$). Through the lens of the structural model, such a gap results from younger generations of households choosing to supply more labor of skill types

that are more complementary to the H technology.

Finally, along the transition, relative employment x_t increases over time as older generations are replaced with younger generations at rate δ , and the relative wage ω_t declines in turn. In the long-run, relative employment and wages are higher than in the initial steady state.

We note that the only source of dynamics in our model is the *differential* labor supply responses of younger and older generations. If households do not respond to changes in the present discounted value of wages ($\psi = 0$) or relative wages do not change ($\Delta \log(\omega) = 0$), there are no transitional dynamics and the responses in the long- and short-run are identical.

Diffusion. One interpretation of the above process is that the technological innovation slowly diffuses in the economy as more workers with the appropriate skills operate the H technology over time. Our model thus formalizes the idea in [Rosenberg \(1972\)](#) that the accumulation of complementary skills shapes the diffusion process. However, this interpretation overlooks the fact that our model does not give rise to an S-curve, which is a common observed feature in practice ([Griliches, 1957](#); [Comin and Mestieri, 2018](#)). Instead, as the figure above shows, relative employment jumps on impact. The reason is that we have assumed that both the innovation increases productivity immediately and existing workers can relocate instantaneously. We have done so for simplicity; to highlight the dynamics induced by changes in relative labor supply over a time frame of generations. Without qualitatively changing any of our main insights, one could “smooth-out” the initial jump of our economy (and generate an S-curve) by enriching our model with any of the well-studied elements that produce a sluggish adoption at relatively shorter horizons, such as the slow increase in productivity due to firm-level learning ([Atkeson and Kehoe, 2007](#)) or the slow reallocation of existing workers ([Matsuyama, 1992](#); [Beraja and Zorzi, 2022](#)).

5 Skill Specificity and Technological Transitions

Through the lens of our model, the evidence in [Section 2](#) can be interpreted as showing that the short-run labor supply elasticity η was lower during the ICT transition in part because of higher skill specificity. Motivated by this, we next analyze the effect of skill specificity on equilibrium dynamics. In particular, [Section 5.2](#) studies the effect of skill specificity on how slow technological transitions are. This is our main outcome of interest for two reasons. First, we wish to speak to the evidence showing a slower adjustment following ICT innovations. Second, as [Section 7](#) shows, the welfare implications substantially differ between fast and slow technological transitions.

Our main theoretical result establishes that a technological transition is slower whenever skill specificity is higher. The reason is that the adjustment is driven more by the entry of new generations of workers instead of the relatively faster reallocation of existing generations that are already present in the economy when the technological innovation arrives.

5.1 The adjustment in high and low η economies

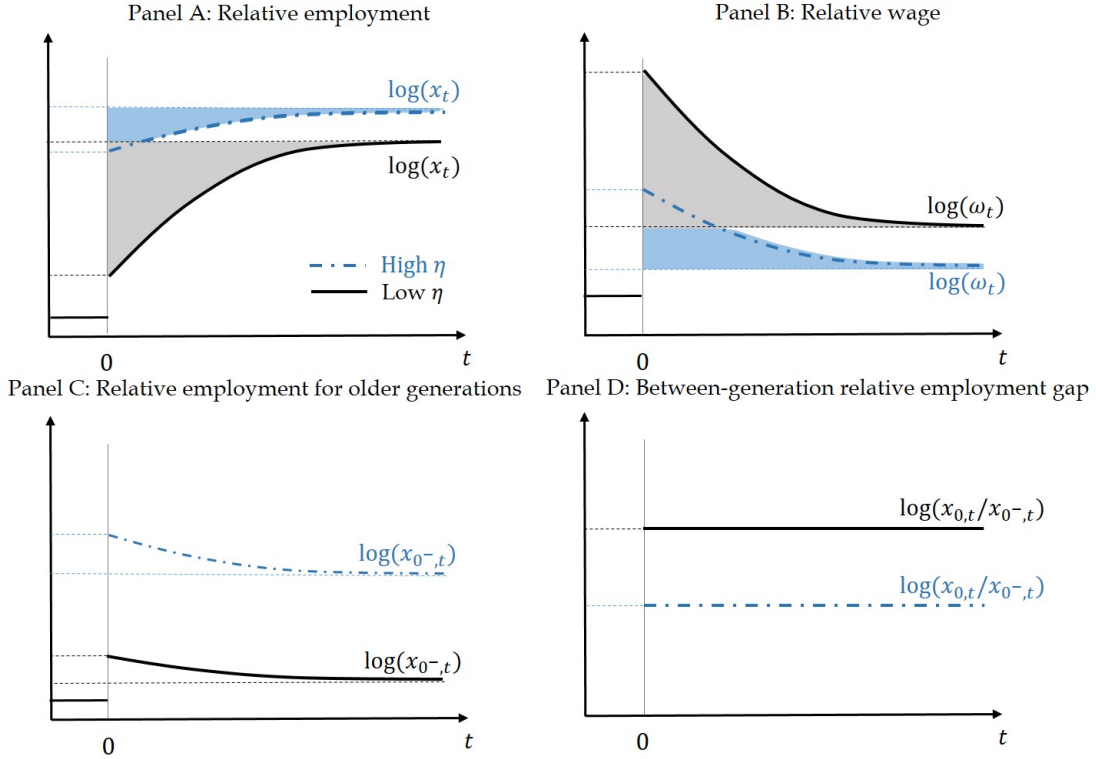
We begin by comparing the adjustment of two economies featuring a high and a low short-run labor supply elasticity η . We interpret this as a comparison between two episodes of adjustment triggered by distinct types of technological innovations that impact different sets of economic activities. In one episode, an innovation augments the productivity of the H technology by $\Delta \log(A)$, where H corresponds to occupations or sectors intensive in skills that are *markedly different* from those used in the rest of the economy (represented by the L technology). That is, this episode is associated with a strong skill specificity economy where the short-run labor supply elasticity η is low. In the other episode, the productivity of the H technology is again augmented by $\Delta \log(A)$ but H instead corresponds to occupations or sectors intensive in *similar* skills to those used in the rest of the economy. That is, this episode is instead associated with a weak skill specificity economy where the short-run labor supply elasticity η is high.

Figure 4 illustrates the adjustment in these two economies. The black solid lines show the responses of an economy with a low elasticity η and the blue dashed lines with a high one. In **Appendix B.3**, we support the graphical representation in **Figure 4** with **Proposition A.1** formally establishing how skill specificity affects the short- and long-run responses, the cumulative impulse response, and the rate of convergence.

In the short-run, when skill specificity is weaker (higher η), older workers of generations born before shock reallocate more ($\Delta \log(x_{0-,0})$ is larger). That is, the within-generation adjustment margin is stronger. As a result, relative wages increase less in the short-run ($\Delta \log(\omega_0)$ is smaller) whereas relative employment increases more ($\Delta \log(x_0)$ is larger). The smaller and less persistent increase in relative wages (and so their present discounted value q) then implies that young entering generations increase their relative labor supply by less compared to older generations alive before the shock ($\Delta \log(x_{0,0}/x_{0-,0})$ is smaller). In other words, the cross-generation adjustment margin becomes *endogenously* weaker. In our structural model, this is because, when q increases by less, younger entering households have weaker incentives to supply labor of skill types that are complementary to the H technology.

Ultimately, transitional dynamics in relative employment x_t and wages ω_t are less important because there are less persistent changes in the relative labor supply that take place as younger generations replace older generations. Formally, we measure

Figure 4: Comparative Statics: Strong and Weak Skill Specificity



this as the cumulative impulse response function being smaller — graphically, the blue shaded areas being smaller than the black shaded areas. The next section shows how the (discounted) cumulative impulse response connects to how slow the adjustment is.

We emphasize that the elasticity η affects cross-generation employment differences *only* because we solve our model in general equilibrium where the relative wage is endogenous. If we had instead taken the relative wage as exogenous, changes in η would impact the labor supply responses of older workers but not the differential responses of younger generations. In this case, η would also not affect transitional dynamics nor how slow the adjustment is.

5.2 How slow is the adjustment?

We conclude this section by showing how different degrees of skill specificity affect our main outcome of interest: how slow technological transitions are. We first define a measure that summarizes the importance of transitional dynamics and, therefore, gives a formal way of quantifying whether the economy's adjustment is fast or whether it slowly plays out over many generations. Specifically, we define the discounted cumulative impulse response (DCIR).

Intuitively, the DCIR is the answer to the question: from the point of view of gen-

erations alive just before the arrival of a technological innovation, how different is the adjustment they expect to see during their lifetime compared to the long-run adjustment? We will say that the economy's adjustment is *slower* when existing generations expect to miss more of the overall adjustment (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 innovation have exponentially distributed death probabilities with rate δ . This is a convenient measure of the importance of transitional dynamics in our context for two reasons. First, it encodes not only the convergence rate 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-run 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 suggests a slow adjustment. 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.

The next proposition shows that, according to this measure, technological transitions are slower whenever skill specificity is stronger. This is because, as [Figure 4](#) illustrated, stronger skill specificity mutes the reallocation of older generations of workers in the short-run *causing* larger endogenous changes in relative wages and, in turn, in the relative labor supply of younger entrant generations.²² As a result, when technological innovations are biased towards economic activities intensive in skills which differ more from those used in the rest of the economy, the adjustment in relative employment is slower since it is driven more by the reallocation across rather within generations.

Proposition 3 (DCIR comparative statics with respect to η) *Following a technological innovation $\Delta \log(A)$, relative employment (x) and wages (ω) adjust slower in economies featuring stronger skill specificity (lower η). Formally,*

$$\frac{\partial DCIR(x)}{\partial \eta} < 0, \quad \frac{\partial DCIR(\omega)}{\partial \eta} < 0.$$

²²We conjecture that this intuition holds 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 relative price changes — e.g., if old vintages of physical capital are less adaptable to new sectors, then the flow of entrants with newer capital vintages will be larger.

Proof. We have that $DCIR(\omega) = \frac{\lambda\delta}{\lambda+\delta} \left| \frac{\int_0^\infty \hat{\omega}_t dt}{\Delta \log(A)} \right|$. From [Proposition A.1](#) in [Appendix B.3](#), we know that λ and $|\int_0^\infty \hat{\omega}_t dt|$ are both smaller when η is larger. Thus, $\frac{\partial DCIR(\omega)}{\partial \eta} < 0$. Since $DCIR(x) = \theta DCIR(\omega)$, this implies $\frac{\partial DCIR(x)}{\partial \eta} < 0$. ■

6 Back to the Evidence

Having established our main theoretical results, we are now ready to come back to the evidence in [Section 2](#) about how the U.S. economy adjusted to ICT and manufacturing innovations. In [Section 6.1](#), we show that the model can match the distinct adjustment paths in both episodes when the short-run elasticity of labor supply is assumed to be lower in the ICT transition. [Section 6.2](#) shows that potential alternative explanations deliver predictions that are odds with the evidence in one or many dimensions. [Section 6.3](#) discusses the role of key model features in accounting for the evidence. [Section 6.4](#) shows that our main results are robust to a number of extensions of our baseline model.

6.1 Model vs. evidence

We map the set of occupations mostly impacted by a technological innovation — i.e., ICT- and manufacturing-intensive occupations — to high-tech production (H) in the model. Likewise, in every time t , we map the group of “younger” workers to those with an age below 29 years in the model (i.e., born between t and $t - 29$) and map the group of “older” workers to those older than 29. Therefore, the relative employment of older generations is $x_t^{older} \equiv \delta \int_{-\infty}^{t-29} e^{\delta(\tau-t)} x_{\tau,t} d\tau$ and that of younger generations is $x_t^{younger} \equiv \delta \int_{t-29}^t e^{\delta(\tau-t)} x_{\tau,t} d\tau$.

[Table 3](#) shows the calibration of the parameters in the reduced-form model. We set the discount rate to $\rho = 0.02$ so that future wages are discounted at an interest rate of 2% per year. We pick a death rate of $\delta = 1/36$ so that young workers aged 29 expect to work until age 65. We set the elasticity of relative labor demand to $\theta = 3$ to match the elasticity of substitution across occupations in [Hsieh et al. \(2019\)](#). We pick a long-run elasticity of relative labor supply of $\psi = 0.4$. In the structural model, this elasticity is interpreted as the elasticity of occupational choices of young workers to changes in the present discounted value of relative earnings. Using experimental variation in beliefs about future earnings in different college majors, [Wiswall and Zafar \(2015\)](#) estimate that a 1% increase in expected relative earnings increases the log odds of majoring in a field by about 0.4% for freshmen and sophomores. Our chosen ψ is in line with this estimate, and is also close to the aggregate labor supply elasticity estimates in [Chetty \(2012\)](#). Finally, for our baseline calibration, we set the short-run relative labor supply elasticity

η to zero, consistent with the lack of relative employment responses for older workers during the ICT episode reported in Table 1. In the structural model, this corresponds to an economy with infinite skill specificity where worker skills cannot be transferred across sectors.

Table 3: Parameter Calibration

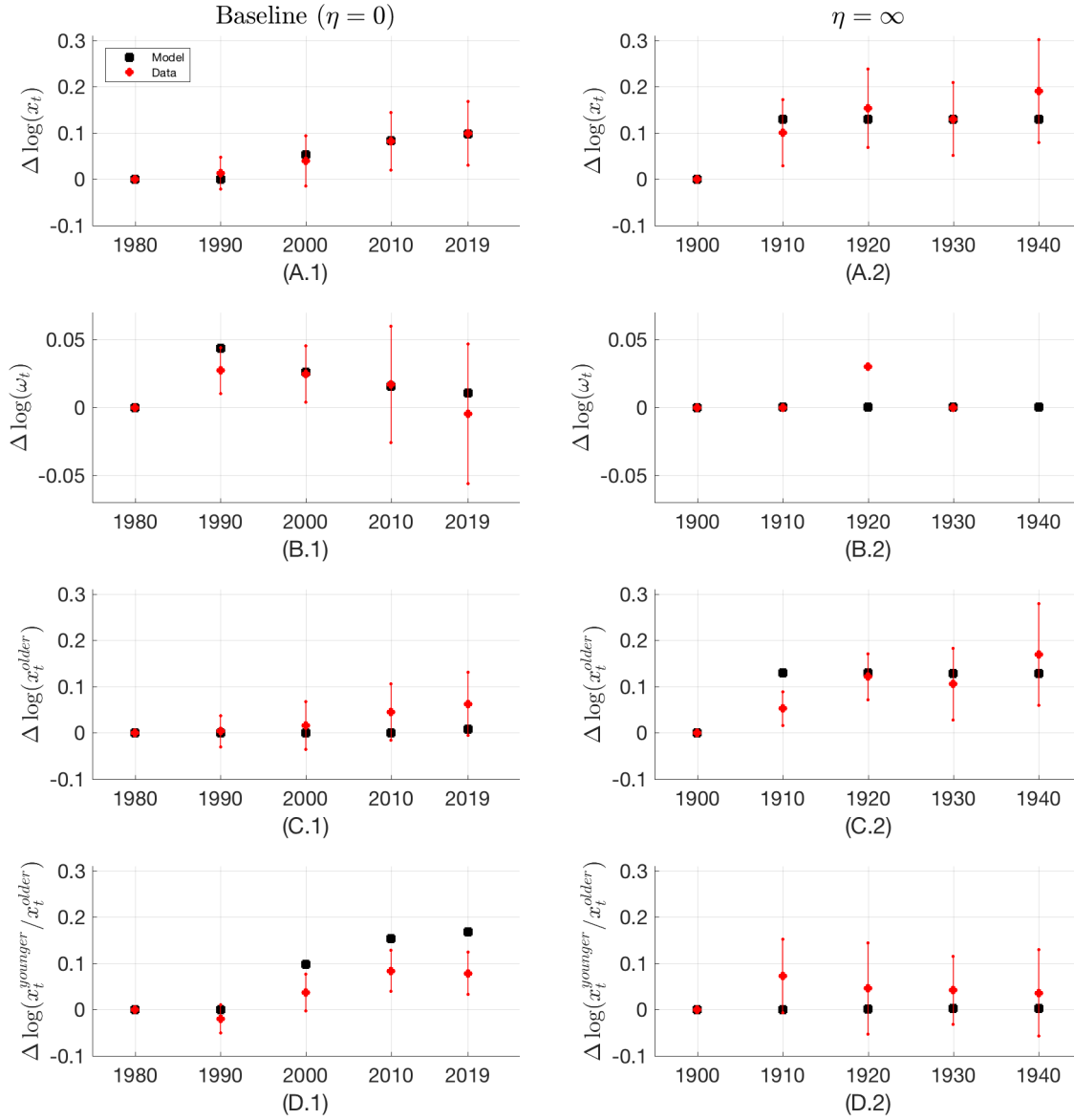
Parameter	Value	Description	Target/Source
ρ	0.02	Discount rate	2% annual interest rate
δ	1/36	Death rate	Workers aged 29 expect to work until 65
θ	3	Labor demand elasticity	Hsieh et al. (2019)
ψ	0.4	Long-run labor supply elasticity	Wiswall and Zafar (2015); Chetty (2012)
η	0	Short-run labor supply elasticity	Table 1

The left column of Figure 5 shows the empirical estimates of the impact of one standard deviation higher ICT exposure on relative outcomes across occupations over time (red diamonds), along with theoretical impulse responses for the same outcomes obtained from our calibrated model (black squares). We set the shock $\Delta \log(A)$ so that the model matches the relative employment response between 1980 and 2019. In the model, we interpret the period between 1980 and 1990 as “the short-run” when the technological innovation arrives. Thus, we associate the year 1990 to the impact responses at $t = 0$ when the shock $\Delta \log(A)$ hits the economy.

The model matches the magnitude and timing of the empirical responses remarkably well. It almost perfectly predicts the slow increase in relative employment (panel A.1) and the inverted-U shape response in relative wages (panel B.1). Crucially, the model achieves this through the right margins of adjustment: it replicates the lack of reallocation of older workers (panel C.1), and the increase in the relative employment of younger workers compared to older ones (panel D.1). The model only slightly overstates the differences in relative employment across generations, and somewhat understates the increase for old generations (although these are not significantly different from zero). Finally, we find a $DCIR(x)$ of approximately 0.5, implying that generations born before 1980 expect to miss half of the overall adjustment because of how slow it was.

The right column of Figure 5 shows the empirical estimates of the impact of one standard deviation higher manufacturing exposure on relative outcomes across occupation between 1900 and 1940 (red diamonds), together with the model-implied impulse responses when the short-run relative labor supply elasticity is large ($\eta \rightarrow \infty$) instead of our baseline zero elasticity (black squares). In the structural model, a higher value of η arises from weaker skill specificity, in line with the lower task distance between the

Figure 5: Dynamic Responses to a Technological Innovation: Model v. Evidence



Note. Red diamonds correspond to the estimates of the impact of one standard deviation higher exposure to ICT in column 1 and manufacturing in column 2 on the log-change of outcomes across occupations — employment in panel A, wages in panel B, employment of older workers in panel C, and the difference in employment of younger and older workers in panel D. Estimates obtained from specification described in Section 2.2, with vertical bars denoting the associated 90% confidence intervals. Black squares are the transitional dynamics predicted by our model following a shock $\Delta \log A$ that we set to match the long-run relative employment response in Panel A.1. Column 1 uses parameters in Table 3, and column 2 the same parameters except that we set $\eta = \infty$.

most and least exposed occupations that we documented in Section 2.3 for manufacturing innovations. The large value that we use corresponds to an economy with weak skill specificity where worker skills are easily transferable across sectors. We use the same values for all other parameters, as well as the magnitude of the shock $\Delta \log A$.

For η large enough, the model predicts a fast increase in relative employment (panel

A.2) coupled with weak changes in relative wages following the technological innovation (panel B.2). Again, the margins of adjustment are consistent with the evidence: the model generates a fast increase in the relative employment of older workers (panel C.2), and similar responses for younger and older generations (panel D.2). The faster transition implies a $DCIR(x)$ of zero, so that generations born prior to the shock expected to see the full adjustment within their lifetime.

Taken together, these results show that a lower short-run elasticity η can quantitatively explain why the adjustment to the ICT innovations of the late twentieth century was slower, more unequal, and mainly driven by the entry of young generations when compared to the adjustment triggered by the innovations in manufacturing of the early twentieth century. We note that the model needs a large change in η to match the large difference in the magnitude of the reallocation of older generations in the two episodes. Our evidence in [Section 2.3](#) shows that skill specificity contributed to reduce the labor supply elasticity of older workers into ICT-intensive occupations. However, we acknowledge that the exercise above does not quantify whether the higher task distance of ICT-intensive occupations (compared to manufacturing-intensive occupations) fully accounts for the lower η during the ICT transition.

6.2 Alternative explanations

We now use our model to evaluate alternative explanations for the differences across the transitions triggered by innovations in ICT and manufacturing. Specifically, we next re-do the analysis above for changes in model parameters that capture other potential reasons for the distinct adjustment patterns observed during the two transitions. We conclude that none of them can jointly explain the differences that we document.

Slow-moving A_t . Technological transitions may be slower or faster due to factors entirely unrelated to specificity and skills. For instance, technologies themselves may only improve gradually and their price decline slowly. Or knowledge about how to use new technologies may be accumulated gradually. In our model, these type of stories would affect productivity and, thus relative labor demand, directly. They can be captured in reduced-form as changes in A_t happening gradually over time as opposed to our baseline of a permanent one-time shock $\Delta \log(A)$. If technologies differ in the speed at which A_t evolves, so will the adjustment. This could in part explain why the ICT and manufacturing transitions differed.²³

However, this explanation is at odds with two facts. First, a slower relative labor demand increase during the ICT transition would have implied a slower increase not

²³For example, ICT could have been held back by the slow rollout of distribution and network technologies that were key complementary inputs.

only of relative employment but also of relative wages. This is not what we observed: the rise in the relative wage of ICT-intensive occupations happened decades before the rise in their relative employment. Second, in an economy without skill specificity ($\eta = 0$) or differences in relative labor supply across generations ($\psi = 0$), slow-moving changes in labor demand by themselves do not lead to differences in employment reallocation across generations. As such, these other sources of slow adjustment dynamics cannot explain the observed differences across younger and older generations that we document. This suggests that mechanisms based on slower dynamics of labor demand cannot alone explain the differences between the two technological transitions.

Changes in ψ . One potential reason behind the different responses in the two episodes is that it was easier for younger workers to acquire the skills required to operate the new technologies used in manufacturing in the early 1900s. That is, the elasticity ψ used to be higher. Indeed, this narrative is in line with historical accounts in [Goldin and Katz \(2009\)](#) and [Alon et al. \(2018\)](#) showing that schooling costs, for example, used to be lower in the early twentieth century than in recent decades.

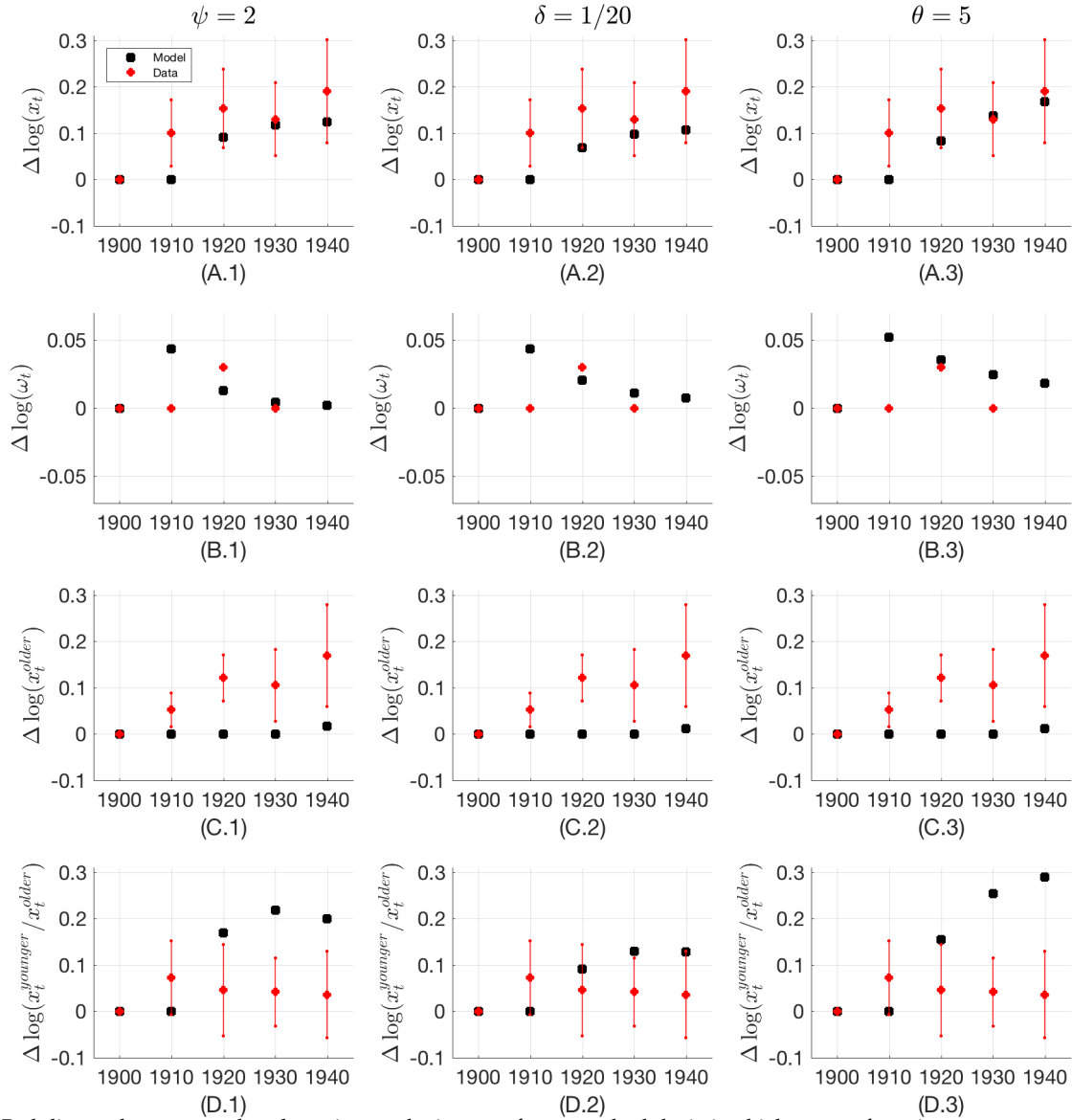
The left column of [Figure 6](#) shows the model-implied responses for $\psi = 2$ instead of our baseline 0.4, again together with the empirical estimates for the impact of higher manufacturing exposure across occupations. A larger ψ cannot explain the different short-run responses in the two episodes: namely, the larger employment response and the smaller wage response over the initial decades of the transition to manufacturing (panels A.1 and B.1). That is, a larger ψ gets the timing of the adjustment wrong. The reason is that this amplifies the relative employment response of young generations (panel D.1), but it cannot generate the strong response observed for older incumbent generations (panel C.1). Thus, any explanation based on the cost of acquiring skills for younger workers is inconsistent with the evidence that employment responses were stronger for both older and younger workers.²⁴

Changes in δ . Another possibility is that, early in the twentieth century, older generations were replaced by younger ones more often. That is, the rate δ used to be higher because of the shorter life expectancy of workers in the early 1900s.

The middle column of [Figure 6](#) shows the model-implied responses for $\delta = 1/20$ instead of our baseline $1/36$. An increase in δ has an effect that is similar to an increase in ψ . Intuitively, both of these parameters govern the elasticity of relative labor at longer horizons. As such, they are unable to simultaneously generate the stronger reallocation

²⁴Note that, if ψ was smaller instead, then the model could generate the small differences across generations that we observe in reallocation towards manufacturing-intensive occupations. However, this would imply a larger and more persistent relative wage increase and a smaller employment response at all horizons; the opposite of what we documented.

Figure 6: Alternative Explanations of Differences across Technological Transitions



Note. Red diamonds correspond to the estimates the impact of one standard deviation higher manufacturing exposure on the log-change of outcomes across occupations – employment in panel A, wages in panel B, employment of older workers in panel C, and the difference in employment of younger and older workers in panel D. Estimates obtained from specification described in Section 2.2, with vertical bars denoting the associated 90% confidence intervals. Black squares are the transitional dynamics predicted by our model following a shock $\Delta \log A$ that we set to match the long-run impact of ICT exposure on relative employment. Predictions obtained with parameters in Table 3, except that $\psi = 2$ in column 1, $\delta = 1/20$ in column 2, and $\theta = 5$ in column 3.

of older workers and the smaller between-generation differences in relative employment responses that we observed during the transition triggered by innovations in manufacturing (panels C.2 and D.2). Consequently, they are also unable to generate the smaller (larger) relative wage (employment) increase that we document at short horizons (panels A.2 and B.2).

Changes in θ . When the short-run labor supply elasticity η is zero but the long-run elasticity ψ is positive, as in our baseline calibration, our model is a variant of [Katz and Murphy \(1992\)](#) with an elastic labor supply, similar to [Caselli \(1999\)](#). In this context, we can interpret the parameter θ as the elasticity of substitution between skilled and unskilled workers. The evidence in [Goldin and Katz \(2009\)](#) suggests that such an elasticity declined over the course of the twentieth century. In the early 1900s, the relevant distinction was between those workers with and without a high school degree, with an elasticity of substitution of 5. In the late 1900s, the relevant distinction was between workers with and without a college degree, with a lower elasticity of 1.6.

The right column of [Figure 6](#) shows the model-implied responses for $\theta = 5$ instead of our baseline 3. An increase in θ can explain the larger relative employment response to the innovations in manufacturing of the early 1900s. However, at odds with the evidence, it also implies an even larger response of relative wages (panel B.3) and, as a result, of relative employment differences across generations (panel D.3). This follows from the usual demand-supply logic that, for the same demand shock $\Delta \log(A)$, the relative wage and employment increases are larger when the relative demand elasticity is larger.²⁵

6.3 The role of key model features

In our model, differences in skill specificity help to explain the observed differences in the economy's adjustment during the two technological transitions. This is because stronger skill specificity reduces the reallocation of older incumbent workers, while it endogenously increases that of younger entrant workers. We next discuss the ingredients of the model that are essential to these conclusions.

Endogenous relative wages. The results above are in contrast to quantitative dynamic Roy models where wages are sometimes taken to be exogenous due to the difficulties in computing the general equilibrium (e.g., [Lee and Wolpin \(2006\)](#)). In particular, changes in skill specificity affect the labor supply responses of older workers but not the differential responses of younger generations when the relative wage ω_t (and thus their present discounted value q_t) is exogenously given.

The intuition follows from the discussion around [Figure 4](#). Cross-generation differences arise only when younger generations' relative labor supply respond to changes in the present discounted value of wages. In a model with endogenous relative wages, a higher η leads to a smaller relative wage increase following the technological innovation $\Delta \log(A)$ (due to the stronger reallocation for older workers), which then implies smaller

²⁵Suppose instead that θ used to be smaller. This could now rationalize the smaller relative wages and cross-generation differences in relative employment at longer horizons. However, in contrast with the evidence, it would also imply a smaller employment response at all horizons and no reallocation of older workers.

cross-generation differences in relative employment and a faster overall employment adjustment. This effect is absent when changes in relative wages are exogenously given, since the incentives of households of entering generations to supply labor of skill types complementary to the H technology would not be affected. Finally, we note that, for the same reason, differences in skill specificity do not affect how slow the adjustment is (as measured by the DCIR) in a model with exogenous changes in inequality.

Endogenous labor supply of young generations. Our results are also distinct from models with exogenous within-generation skill heterogeneity — as in static Roy models, e.g., [Costinot and Vogel \(2010\)](#) and [Burstein et al. \(2019\)](#). Our model features a fixed labor supply of young generations when the long-run labor supply elasticity ψ is zero. This rules out the type of cross-generational employment differences implied by ICT innovations. In our model, it also implies that all responses are instantaneous (i.e., there are no transitional dynamics).

Heterogenous skills and specificity. A special case of our model without transferable skills — i.e., infinite skill specificity, $\eta = 0$ — implies that old generations do not reallocate after a technological innovation, which is inconsistent with the evidence in [Section 2.2](#) for the manufacturing innovations of the early 1900s. This would also be the case in models without worker mobility (e.g., [Guerreiro et al. \(2017\)](#)) or where the innovation is biased towards certain observable worker attributes like education or age (e.g., [Katz and Murphy \(1992\)](#), [Caselli \(1999\)](#) or [Card and Lemieux \(2001\)](#)) instead of impacting economic activities using specific skills.

At the other extreme, we can consider a model where skills are perfectly transferable — i.e., no skill specificity $\eta = \infty$. This case is similar to canonical structural transformation models where labor is homogenous, reviewed by [Herrendorf et al. \(2014\)](#). It implies that there are no cross-generational employment differences, which is not consistent with our evidence for the ICT episode. This would be the case as well in models where the dynamics of (homogeneous) labor reallocation are driven by shifters of labor demand, like firms investment and entry decisions (e.g., [Atkeson and Kehoe \(2007\)](#), [Dix-Carneiro and Kovak \(2017\)](#)). Note also that this class of models cannot explain our findings that, upon arrival in the 1980s, ICT innovations triggered a fast increase in relative wages that preceded relative employment responses by decades.

6.4 Additional determinants of adjustment dynamics

The theory so far has ignored several determinants of adjustment dynamics. In [Appendix C.2](#), we present three extensions that relax some of the assumptions of our baseline model. For all extensions, our comparative static results with respect to η in

Propositions 3 and the numerical results in this section remain valid. However, as we discuss below, the extensions affect the *levels* of the DCIRs and the relative employment responses for older and younger generations.

Learning-from-others. Our first extension considers a "learning-from-others" externality. Specifically, we relax the assumption that the labor supply choices of young generations depends only on the future path of wages. Instead, we assume that certain skill-types may be easier to supply because workers can "learn from others" when those types are already abundant in the economy. This extension introduces a backward-looking element into the household's problem and complementarities in the labor supply decisions across generations. This is similar to the cross-generation complementarities that arise in the environment considered by Chari and Hopenhayn (1991). In this case, dynamic responses are qualitatively similar to those of our baseline economy when ψ is higher and δ is lower. This makes the adjustment slower, the relative employment response of older workers at long horizons smaller, and the cross-generation differences in relative employment at long horizons larger.

Re-training of old workers. Our second extension relaxes the assumptions that households choose their supply of skill types only at birth. We allow an exogenous fraction of older generations that were present before the technological innovation to re-optimize their labor supply *as if* they were a young generation entering at time $t = 0$. One interpretation of this is that older workers can now re-train after the innovation arrives. This extension yields responses that are qualitatively similar to our baseline when the short-run labor supply elasticity η is higher. Thus, it makes the adjustment faster, the employment responses of older generations stronger, and the cross-generation differences in employment lower.

Population growth. Our third extension allows for the birth rate to be higher than the death rate. This raises the convergence rate λ , resulting in a faster adjustment.

7 Welfare Implications of Skill Specificity

So far, we have shown that the model matches in a parsimonious way the evidence on *relative* responses across occupations and generations for the transitions triggered by innovations in ICT and manufacturing. This gives us confidence that we can use the model to study the *aggregate* implications of technological innovations. In particular, we next analyze how welfare is affected by the economy's degree of skill specificity. This normative analysis requires us to move away from the reduced-form model of Section 4 and commit to the structural model of Section 3.

To obtain an average welfare measure across generations W , we take an utilitarian approach by considering a population-weighted average of the utility of different generations, where generation- τ 's weight is $\delta\rho e^{-\rho\tau}$ as in [Calvo and Obstfeld \(1988\)](#) and their utility U_τ is defined in (11). This implies that aggregate welfare in consumption-equivalent units is

$$W = \rho\delta \int_{-\infty}^{\infty} e^{-\rho\tau} U_\tau d\tau. \quad (20)$$

We consider a second order approximation of the welfare function around the stationary equilibrium. As usual, this higher-order approximation allows us to capture households' preference for smoother consumption paths.²⁶ For the relevant case of a small discount rate ρ , the following proposition shows that the welfare gains from technological innovations are smaller when skill specificity is stronger because the adjustment is slower.

Proposition 4 (Welfare gains) *Let ΔW be the change in welfare from a technological innovation $\Delta \log(A)$ and assume that $\theta > 1$. Then,*

$$\lim_{\rho \rightarrow 0} \frac{\partial \Delta W}{\partial \eta} > 0$$

Proof. See [Appendix B.4](#).

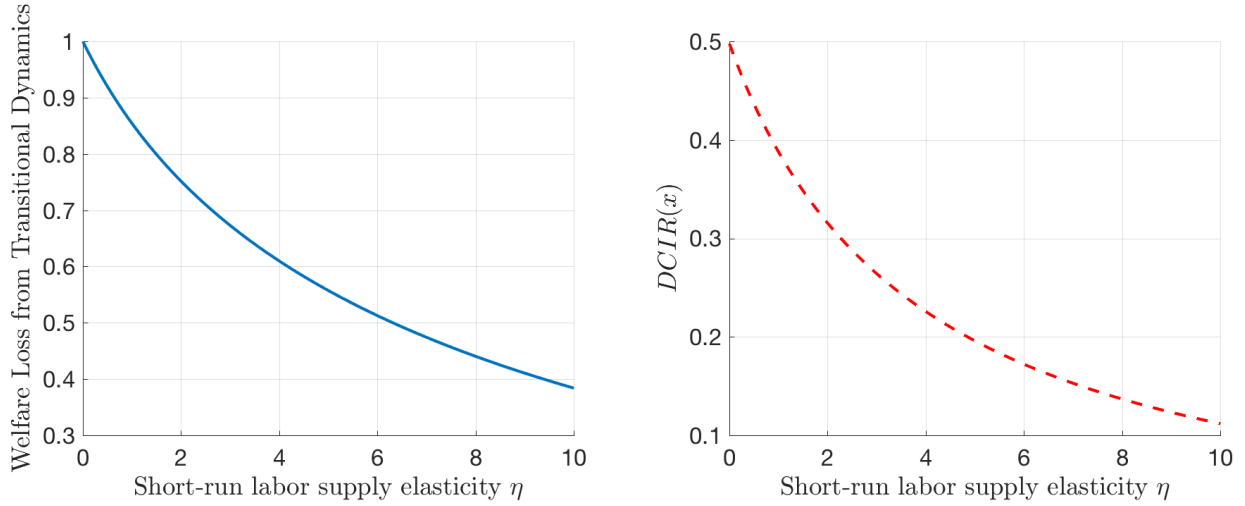
To get a sense of magnitudes and provide intuition for the proposition, [Figure 7](#) uses our calibrated model to compute welfare and the $DCIR(x)$ under different values of the elasticity η . This also allows us to show, numerically, that the result in the proposition still holds for reasonable values of ρ away from the limit $\rho \rightarrow 0$.

The left panel of [Figure 7](#) shows how the elasticity η affects the welfare loss from transitional dynamics. We compute this loss as $\Delta W - \Delta U_\infty$. That is, the difference between the actual welfare gains and the gains associated with the long-run change in utility between the initial and final steady states. We note that such long-run gains ΔU_∞ also correspond to an economy where the adjustment is instantaneous, such as when the elasticity η is infinity. Moreover, to ease interpretation, we normalize the loss to 1 for the economy from our baseline calibration with infinite specificity ($\eta = 0$). The right panel of the figure shows the $DCIR(x)$ as a measure of how slow the adjustment is.

For our baseline calibration with $\eta = 0$, the $DCIR(x)$ is 0.5. As η increases, the adjustment becomes faster and the welfare loss decreases. For example, when η increases to 3, the $DCIR(x)$ is cut in half to about 0.25. In turn, this faster adjustment implies that the welfare loss from transitional dynamics are about 1/3 smaller compared to the baseline. Underlying this welfare calculation are significant consumption differences

²⁶For example, second order approximations are common in the business cycles literature when studying the welfare losses from fluctuations in output and inflation ([Woodford, 2011](#)).

Figure 7: Welfare in Fast and Slow Technological Transitions



across generations. Early generations have much smaller consumption than later ones which entered the market when average wages were higher and the relative wage ω lower. Such differences across generations decrease as η increases because there is a faster and stronger increase in average wages and a less persistent and smaller increase in relative wages.

8 Conclusions

Technological transitions are not always alike. Labor markets adjust rapidly to some technological innovations and more slowly to others. One important reason is that certain innovations benefit production activities that require more specific skills. Following such innovations, the adjustment is slower because it relies more on the gradual entry of young generations of workers to the benefitted activities and less by the faster reallocation of older incumbent workers. To support this argument, we first present evidence on how U.S. labor markets adjusted to manufacturing innovations in the early twentieth century and ICT later in that century. We then build a model of technological transitions that allows us to sharply characterize the impact of skill specificity on equilibrium dynamics, interpret our evidence, and analyze normative implications.

The evidence shows that the adjustment to ICT innovations was slower and more unequal. Differences in the reallocation of younger and older generations account for the slower expansion of ICT-intensive occupations. While the reallocation to manufacturing occupations in the early twentieth was similar for young and old generations, older generations did not move towards ICT-intensive occupations. Instead, the ICT transition was entirely driven by the gradual entry of younger generations. Stronger skill specificity during the ICT transition can partly explain these facts.

Going forward, our results raise the question of whether recent advances in Artificial Intelligence (AI) — a new general-purpose technology that could transform multiple sectors and occupations — resemble past innovations in ICT or manufacturing. Will the adjustment be slow, mainly driven by the entry of younger generations, or will it be fast, as the skills possessed by older incumbent generations can easily be transferred to work in activities benefiting from AI?

References

- Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4, pages 1043–1171. Elsevier.
- Acemoglu, D. and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542.
- Adão, R., Beraja, M., and Pandalai-Nayar, N. (2020). Technological transitions with skill heterogeneity across generations. Technical Report No. w26625, National Bureau of Economic Research.
- Alon, T. et al. (2018). Earning more by doing less: Human capital specialization and the college wage premium. *manuscript, University of California at San Diego*.
- Atkeson, A. and Kehoe, P. J. (2007). Modeling the Transition to a New Economy: Lessons from Two Technological Revolutions. *American Economic Review*, 97(1):64–88.
- Autor, D. and Dorn, D. (2009). This job is "getting old": Measuring changes in job opportunities using occupational age structure. *American Economic Review*, 99(2):45–51.
- Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, 103(5):1553–97.
- Autor, D. H., Katz, L. F., and Krueger, A. B. (1998). Computing inequality: have computers changed the labor market? *The Quarterly journal of economics*, 113(4):1169–1213.
- Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.
- Beraja, M. and Zorzi, N. (2022). Inefficient automation. Technical report, MIT mimeo.

- Buera, F. J., Kaboski, J. P., Mestieri, M., and O'Connor, D. G. (2020). The stable transformation path. Technical report, National Bureau of Economic Research.
- Buera, F. J., Kaboski, J. P., Rogerson, R., and Vizcaino, J. I. (2021). Skill-biased structural change.
- Burstein, A., Morales, E., and Vogel, J. (2019). Changes in between-group inequality: computers, occupations, and international trade. *American Economic Journal: Macroeconomics*, 11(2):348–400.
- Calvo, G. A. and Obstfeld, M. (1988). Optimal time-consistent fiscal policy with finite lifetimes. *Econometrica: Journal of the Econometric Society*, pages 411–432.
- Card, D. (2005). Is the new immigration really so bad? *The economic journal*, 115(507):F300–F323.
- Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.
- Card, D. and Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis. *The Quarterly Journal of Economics*, 116(2):705–746.
- Caselli, F. (1999). Technological revolutions. *American economic review*, 89(1):78–102.
- Chari, V. V. and Hopenhayn, H. (1991). Vintage human capital, growth, and the diffusion of new technology. *Journal of political Economy*, 99(6):1142–1165.
- Chetty, R. (2008). Moral hazard versus liquidity and optimal unemployment insurance. *Journal of political Economy*, 116(2):173–234.
- Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018.
- Comin, D. and Mestieri, M. (2018). If technology has arrived everywhere, why has income diverged? *American Economic Journal: Macroeconomics*, 10(3):137–78.
- Costinot, A. and Vogel, J. (2010). Matching and inequality in the world economy. *Journal of Political Economy*, 118(4):747–786.
- David, P. A. (1990). The dynamo and the computer: an historical perspective on the modern productivity paradox. *The American Economic Review*, 80(2):355–361.
- Dix-Carneiro, R. (2014). Trade liberalization and labor market dynamics. *Econometrica*, 82(3):825–885.

- Dix-Carneiro, R. and Kovak, B. K. (2017). Trade liberalization and regional dynamics. *American Economic Review*, 107(10):2908–46.
- Dvorkin, M. and Monge-Naranjo, A. (2019). Occupation mobility, human capital and the aggregate consequences of task-biased innovations. Technical report.
- Eckardt, D. (2019). Are chemists good bankers? returns to the match between training and occupation. Technical report, LSE.
- Falck, O., Gold, R., and Heblich, S. (2014). E-lections: Voting behavior and the internet. *American Economic Review*, 104(7):2238–65.
- Gathmann, C. and Schönberg, U. (2010). How general is human capital? a task-based approach. *Journal of Labor Economics*, 28(1):1–49.
- Geroski, P. A. (2000). Models of technology diffusion. *Research policy*, 29(4-5):603–625.
- Goldin, C. and Katz, L. F. (1998). The origins of technology-skill complementarity. *The Quarterly journal of economics*, 113(3):693–732.
- Goldin, C. and Margo, R. A. (1992). The great compression: The wage structure in the united states at mid-century. *The Quarterly Journal of Economics*, 107(1):1–34.
- Goldin, C. D. and Katz, L. F. (2009). The Future of Inequality: The Other Reason Education Matters So Much. Scholarly articles, Harvard University Department of Economics.
- Goos, M., Manning, A., and Salomons, A. (2009). Job polarization in europe. *American economic review*, 99(2):58–63.
- Gordon, R. J. (2000). Does the "new economy" measure up to the great inventions of the past? *Journal of economic perspectives*, 14(4):49–74.
- Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica, Journal of the Econometric Society*, pages 501–522.
- Guerreiro, J., Rebelo, S., and Teles, P. (2017). Should robots be taxed? Technical report, National Bureau of Economic Research.
- Hayashi, F. (1982). Tobin's marginal q and average q : A neoclassical interpretation. *Econometrica: Journal of the Econometric Society*, pages 213–224.
- Heckman, J. J., Lochner, L., and Taber, C. (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.

- Helpman, E. (1998). *General purpose technologies and economic growth*. MIT press.
- Herrendorf, B., Rogerson, R., and Valentinyi, Á. (2014). Growth and structural transformation. In *Handbook of economic growth*, volume 2, pages 855–941. Elsevier.
- Hobijn, B., Schoellman, T., and Vindas, A. (2019). Wages during structural transformation: The importance of cohort labor supply decisions.
- Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and us economic growth. *Econometrica*, 87(5):1439–1474.
- Humlum, A. (2019). Robot adoption and labor market dynamics. *Princeton University*.
- Jovanovic, B. and Rousseau, P. L. (2005). General purpose technologies. In *Handbook of economic growth*, volume 1, pages 1181–1224. Elsevier.
- Kambourov, G. and Manovskii, I. (2008). Rising occupational and industry mobility in the united states: 1968–97. *International Economic Review*, 49(1):41–79.
- Katz, L. F. and Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *The quarterly journal of economics*, 107(1):35–78.
- Lagakos, D., Moll, B., Porzio, T., Qian, N., and Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, 126(2):797–849.
- Lee, D. and Wolpin, K. I. (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica*, 74(1):1–46.
- Lucas, R. E. and Moll, B. (2014). Knowledge growth and the allocation of time. *Journal of Political Economy*, 122(1):1–51.
- Macaluso, C. (2022). Skill remoteness and post-layoff labor market outcomes.
- Mansfield, E. (1961). Technical change and the rate of imitation. *Econometrica: Journal of the Econometric Society*, pages 741–766.
- Martellini, P. and Menzio, G. (2020). Declining search frictions, unemployment, and growth. *Journal of Political Economy*, 128(12):4387–4437.
- Matsuyama, K. (1992). A simple model of sectoral adjustment. *The Review of Economic Studies*, 59(2):375–387.
- Michaels, G., Rauch, F., and Redding, S. J. (2012). Urbanization and structural transformation. *The Quarterly Journal of Economics*, 127(2):535–586.
- Murphy, K. M. and Topel, R. H. (1987). The evolution of unemployment in the united states: 1968-1985. *NBER macroeconomics annual*, 2:11–58.

- Perla, J. and Tonetti, C. (2014). Equilibrium imitation and growth. *Journal of Political Economy*, 122(1):52–76.
- Piketty, T. and Saez, E. (2003). Income inequality in the united states, 1913–1998. *The Quarterly journal of economics*, 118(1):1–41.
- Porzio, T., Rossi, F., and Santangelo, G. (2020). The Human Side of Structural Transformation.
- Rosenberg, N. (1972). Factors affecting the diffusion of technology. *Explorations in economic history*, 10(1):3.
- Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, 3(2):135–146.
- Sanderson, E. and Windmeijer, F. (2016). A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221.
- Schubert, G., Stansbury, A., and Taska, B. (2021). Employer concentration and outside options. *Available at SSRN 3599454*.
- Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of labor economics*, 24(2):235–270.
- Traiberman, S. (2019). Occupations and import competition: Evidence from denmark. *American Economic Review*, 109(12):4260–4301.
- Violante, G. L. (2002). Technological acceleration, skill transferability, and the rise in residual inequality. *The Quarterly Journal of Economics*, 117(1):297–338.
- Wiswall, M. and Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, 82(2):791–824.
- Woodford, M. (2011). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Appendix for: Fast and Slow Technological Transitions

This appendix contains supplemental material for the article “Fast and Slow Technological Transitions.” We provide (i) further details on data construction and the evidence, (ii) the proofs of all lemmas and propositions in the paper, and (iii) some additional theoretical results.

Any references to equations, figures, tables, assumptions, propositions, lemmas, or sections that are not preceded “A.” – “C.” refer to the main article.

Appendix A Evidence from Two Technological Transitions: Data Construction and Additional Results

This section complements the results presented in [Section 2](#). We provide details about data construction and present additional results.

A.1 Data Construction

This section outlines in detail the construction of the data used in [Section 2](#).

Employment and average wage by occupation and age group. Our main data source is the historical data of the U.S. Census from IPUMS. For each Census year, we select the sample of males in the labor force aged 16-64 years old that have valid occupation information. Following the guidelines provided by IPUMS, we use the 1950 occupation classification before 1970, and the 1990 occupation classification after 1980. For the latter period, we further aggregate occupations into the classification used by [Autor and Dorn \(2013\)](#). We use sampling weights to compute outcomes by occupation and age group. Employment is the sum of the number of individuals in each occupation and group. The average wage is the aggregate annual income (“incwage”) of full-time workers in an occupation-group divided the total number of such individuals. We define full-time workers as those working at least 40 weeks in a year and 35 hours in a week.

Occupation exposure to manufacturing industries. We use the U.S. Census classification to group industries into a manufacturing sector and a non-manufacturing sector. The occupation’s exposure is $\text{Exposure}_o \equiv \sum_k \ell_{o,t_0}^k S^k$, where S^k is a dummy that equals one if industry i is classified as manufacturing (codes 300-500 in the 1950 industry classification), and ℓ_{o,t_0}^k is the share of industry k in the number of employed individuals in occupation o in the 1900 Census.

Occupation exposure to electricity-intensive industries. To construct an occupation’s electricity exposure, we begin by obtaining industry electricity exposure from the 1947 BEA input-output tables. These tables are the oldest available, and are coded in 1957 SIC codes. Our occupation-industry data from IPUMS uses Census industry codes. A concordance between any vintage of Census industry codes and 1957 SIC codes is not available, so we construct a concordance for this match. For this concordance, we map each 2-digit SIC code from the IO tables to corresponding Census industry codes. In the few cases of multiple SIC codes matching to a single Census 1950 industry code, we assign the simple average of the SIC electricity exposure to the Census industry code.

The SIC industry’s electricity exposure is defined as the sum of its input spending (as a share of its output) on “Purchases of Electrical Industrial Equipment and Apparatus”, “Purchases of Electric lighting and wiring”, “Purchases of Electric components and accessories” and “Purchases of Misc. electrical machinery, equipment and supplies.” As the 1947 IO tables do not separate purchases of electricity from other utilities such as gas, water and sanitary services, we do not include direct electricity purchases in our exposure measure. Finally, to mitigate noise in the exposure measure, we then average the exposure measure of each Census 1950 industry in four bins – non-manufacturing, low electricity usage manufacturing, moderate electricity usage manufacturing, and high electricity usage manufacturing. The bins within the manufacturing sector are constructed as the bottom, middle and top 33% of manufacturing electricity exposure. Finally, we compute the occupation’s exposure to electricity as $\text{Exposure}_o \equiv \sum_i \ell_{o,t_0}^k S^k$, where S^k is the electricity input exposure for industry k (computed as described above), and ℓ_{o,t_0}^k is the share of industry k in the number of employed individuals in occupation o in the 1900 Census.

Occupation exposure to ICT-intensive industries. For industry ICT exposure, we first construct for each NAICS industry its spending (as a share of its output) on the following items related to ICT in the 1997 BEA input-output table: “Audio, visual and communications equipment”, “Semiconductors and electronic components”, “Software”, “Motion pictures and sound recordings”, “Radio and television broadcasting”, “Cable networks and program distribution”, “Telecommunications”, “Information services” and “Data processing services.” We use the crosswalk provided by the Census Bureau to map the NAICS 1997 codes to the 1990 Census industry classification. For each Census industry, we use the simple average of the ICT exposures of all associated NAICS industries.

We compute the occupation’s exposure following the same steps described above for the construction of the exposure to electricity-intensive industries. We first group industries into four bins of ICT exposure based on quartiles of industry spending shares on ICT, and use the simple average in each bin as the exposure measure for the Census industries in that bin. We then compute the occupation’s exposure to ICT as $\text{Exposure}_o \equiv \sum_i \ell_{o,t_0}^k S^k$, where S^k is the ICT input exposure for industry k (computed as described above), and ℓ_{o,t_0}^k the share of industry k in the number of full-time employed individuals in occupation o in the 1980 Census.

Occupation task requirements and task-based exposure. We use the replication data from [Autor and Dorn \(2013\)](#) to measure the intensity of each occupation (in the 1990 Census classification) on three types of tasks: manual, routine, and abstract/cognitive. [Autor and Dorn \(2013\)](#) follow the procedure in [Autor et al. \(2003\)](#) to measure task intensity us-

ing the job task requirements published in the fourth edition of the U.S. Department of Labor’s Dictionary of Occupational Titles (DOT) (U.S. Department of Labor 1977) – for details see Section II of [Autor and Dorn \(2013\)](#). For each occupation, the intensity index in a task is a number that varies between 0 and 10. We define $Exposure_o$ as occupation o ’s standardized intensity index in abstract activities.

Occupation education-based exposure. We also build an index of the education requirements of each occupation as the share of college graduates among full-time employed individuals in occupation o in the 1980 Census.

A.2 Additional Empirical Results: Exposure to new technologies and occupational outcomes

[Figure A1](#) investigates the robustness of the patterns documented in [Figure 1](#) of [Section 2.1](#). Panels A and B report respectively the responses of relative employment and wages to ICT exposure in the late twentieth century. Panel C reports the response of relative employment to manufacturing exposure in the early twentieth century. For ease of comparison, the first column in all panels replicates our baseline estimates in [Figure 1](#).

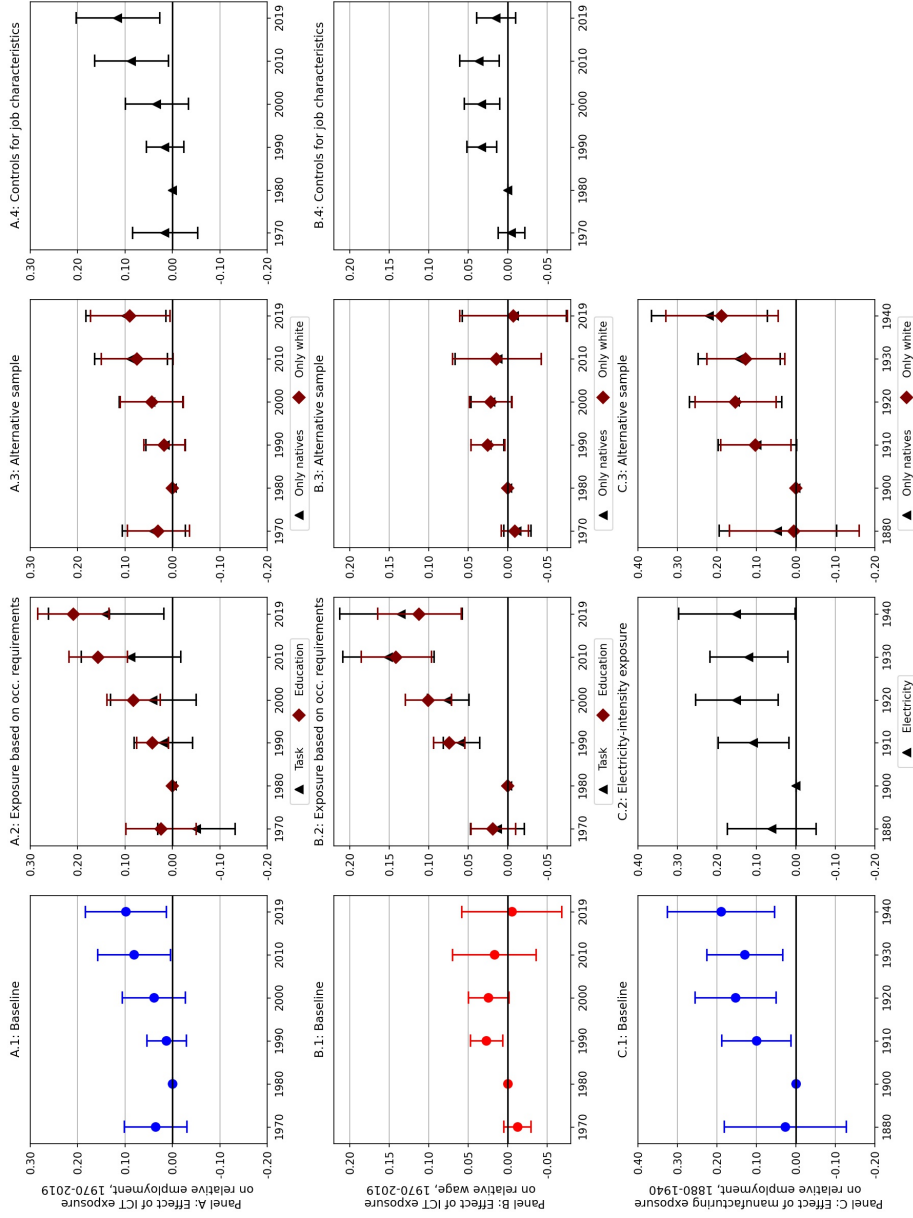
Our first set of robustness checks use alternative exposure measures for the two episodes. For the ICT episode, the extensive literature reviewed in [Acemoglu and Autor \(2011\)](#) has suggested that technology-driven labor demand shocks after 1980 might have increased the productivity of jobs intensive in cognitive tasks (relative to those intensive in routine tasks) or those requiring a college degree (relative to those not requiring higher education). Thus, as described above, we construct alternative exposure measures based on each occupation’s cognitive task intensity and college requirement. For the early episode, motivated by the link between the spread of electricity and the increase in manufacturing in this period ([David, 1990](#); [Jovanovic and Rousseau, 2005](#)), we construct a measure of an occupation’s exposure to electricity usage as detailed above. Column 2 presents estimates obtained with these alternative exposure measures. Overall, results attest the robustness of the conclusions in [Section 2.1](#). Estimates of the impact on relative employment in both episodes are similar, with the education-based exposure yielding a slightly stronger relative employment response after 2000 (although the difference between these estimates and our baseline are not statistically significant at usual levels). All specifications also show a significant increase in the relative wage of more exposed occupations shortly after 1980, while employment responses were small and not statistically different from zero. However, both alternative measures lead to larger and more persistent increases in relative wages after 2000, with the slow down only starting in 2010. We note that, as in our baseline, estimates of relative wage responses after 2000

also have somewhat wide confidence intervals for these alternative exposure measures.

Our second set of robustness checks in column 3 address concerns that two sources of aggregate labor supply shocks during these episodes had heterogeneous effects across occupations that could have been correlated with our exposure measures. Both the early and the late twentieth centuries were periods in which immigration changed substantially. To the extent that immigrants differ from natives in their occupation allocation (Card, 2005), changes in the stock of immigrants may have affected relative employment in more exposed occupations. Column 3 shows that this was not the case, since our estimates are similar when we consider only a subset of U.S. natives. Both periods were also marked by changes in social norms that affected the employment allocation of minorities across occupations (Hsieh et al., 2019). We show that this does not affect our conclusions as estimates are similar when we restrict our sample to include only white males.

Finally, our last set of robustness checks addresses concerns that our estimates could be biased by job amenities that varied across occupations in a manner correlated with their exposure. Unfortunately, the U.S. Census did not collect much information on job attributes in 1900-1940, so we only report estimates with additional controls for the more recent ICT episode. Specifically, we estimate equation (1) controlling for the initial hours worked in an occupation, the self employment rate of the occupation, and the share of full-time workers in the occupation. Column 4 shows that our conclusions are robust to the inclusion of these controls. In fact, they lead to more precise estimates of relative wage responses in ICT-intensive occupations after 2000.

Figure A1: Exposure to Technological Innovations and Occupation Outcomes in the Two Episodes – Robustness



Note. Dots are the point estimates of α_t obtained from (1) in a sample of 310 occupations for Panels A-B and 201 occupations for Panel C with positive employment in the initial and final years of the period for the U.S. Census sample of males aged 16-64yrs in the labor force. Estimates are weighted by the occupation's employment at t_0 . Vertical bars denote 95% confidence intervals clustered by occupation. The occupation exposure in Panels A-B is the occupation's average ICT-intensity across industries, except in Panels A.2 and B.2 where it is either the occupation cognitive intensity or college share, and in Panel C is the occupation's average manufacturing-intensity across industries, except in C.2 where it is the occupation's average electricity-intensity across industries. Exposure measures are normalized to have a unit standard deviation. Column 3 further restricts the samples used to contain only U.S. native males or white males. Column 4 includes additional controls in the regression, as detailed in the text.

A.3 Additional Empirical Results: Responses in occupation outcomes across worker generations

We now provide additional results that complement the estimates in [Section 2.2](#).

Robustness of responses in occupation outcomes for different worker generations. [Table A1](#) investigates the robustness of the patterns in [Table 1](#) to alternative time periods, exposure measures, and sample definitions.

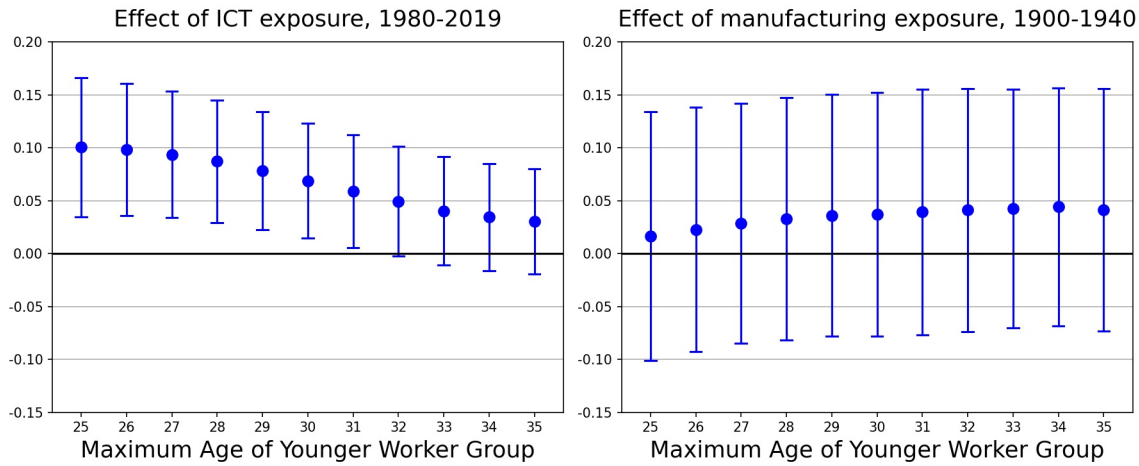
Panel A reports estimates of the effect of ICT exposure on relative employment. For our baseline specification, we find that, in line with estimates in [Figure 1](#), relative employment responses before 2000 were small and non-statistically different from zero at usual levels for both older and younger generations. Note that the employment response for younger generations was already stronger than that for older ones, but the difference reported in column (4) was again small and non-significant. Panel A also shows that the differential employment response between younger and older workers is even stronger than the baseline when we use the alternative measures of ICT exposure – that is, the occupation cognitive intensity or its education requirement. Finally, restricting the sample to native workers or white workers does not change our conclusions.

Panel B reports estimates of the effect of ICT exposure on relative wages. It shows that the lack of differential wage responses remains true at shorter horizons and with alternative samples. When we use the two alternative exposure measures, we find that younger workers actually had *lower* relative wage growth in more exposed occupations. These estimates reinforce our conclusions that relative demand did not increase more for younger workers than older ones conditional on the occupation’s exposure.

Panel C reports estimates of the effect of manufacturing exposure on relative employment in the early twentieth century. In line with our baseline estimates, we find that, compared to the ICT episode, employment growth in manufacturing-intensive occupations in the early 1900s was faster and stronger across both age groups, independent of the specification of the time horizon, sample, or exposure measure. In all specifications, we find that the difference in responses for younger and older generations was small and non-significant.

We next investigate the robustness of the between-generation gap in employment responses in [Table 1](#) with respect to the definition of the two generation groups. We do so by estimating (2) with alternative age cutoffs to define the groups of younger and older workers. [Figure A2](#) reports our estimates of $\beta_{\text{younger},t} - \beta_{\text{older},t}$ for different age cutoffs defining the younger worker group. For the ICT episode, the estimated coefficient becomes smaller as we expand the young group to include older workers. For the transition to manufacturing in 1900-1940, the left panel shows that, if anything,

Figure A2: Differential Reallocation of Younger Workers to More Exposed Occupations in the Two Episodes – Varying Age Thresholds



Note. Dots depict $\beta_{\text{younger},t} - \beta_{\text{older},t}$ obtained from the estimation of equation (2), with the dependent variable being the change in employment over the entire period (1980-2019 in the left panel and 1900-1940 in the right panel). The figures vary the age threshold for defining younger workers. Estimates are weighted by the occupation's employment at t_0 . The occupation exposure in the left panel is the occupation's average ICT-intensity across industries and in the right panel is the occupation's average manufacturing-intensity across industries. Exposure measures are normalized to have a unit standard deviation. Confidence intervals are 95% confidence intervals, computed from the regression where standard errors were clustered by occupation.

the differential reallocation of younger workers becomes stronger as we increase the age cutoff.

Table A1: Exposure to Technological Innovations and Occupation Outcomes for Younger and Older Workers in the Two Episodes – Robustness

Worker group:	All	Older	Younger	Younger -
	16-64yrs	30-64yrs	16-29yrs	Older
	(1)	(2)	(3)	(4)
<i>Panel A: Effect of ICT exposure on relative employment</i>				
Baseline:				
Full period, 1980-2019	0.099** (0.044)	0.063 (0.043)	0.141*** (0.050)	0.078*** (0.029)
Mid period, 1980-2000	0.040 (0.034)	0.015 (0.032)	0.052 (0.044)	0.037 (0.025)
Alternative Exposure Measure:				
Cognitive task requirement	0.141** (0.062)	0.054 (0.064)	0.163** (0.066)	0.109*** (0.031)
Education requirement	0.209*** (0.039)	0.132*** (0.038)	0.254*** (0.047)	0.122*** (0.033)
Alternative Sample:				
Only U.S. natives	0.099** (0.043)	0.067 (0.042)	0.134*** (0.0490)	0.067** (0.027)
Only white	0.090** (0.043)	0.050 (0.042)	0.131*** (0.048)	0.082*** (0.028)
<i>Panel B: Effect of ICT exposure on relative wage</i>				
Baseline:				
Full period, 1980-2019	-0.005 (0.032)	-0.003 (0.035)	0.006 (0.028)	0.009 (0.018)
Mid period, 1980-2000	0.025* (0.013)	0.021 (0.014)	0.039*** (0.012)	0.019 (0.012)
Alternative Exposure Measure:				
Cognitive task requirement	0.136*** (0.040)	0.162*** (0.042)	0.101*** (0.034)	-0.061*** (0.020)
Education requirement	0.112*** (0.027)	0.133*** (0.028)	0.075*** (0.021)	-0.058*** (0.017)
Alternative Sample:				
Only U.S. natives	-0.008 (0.034)	-0.009 (0.036)	0.006 (0.030)	0.015 (0.018)
Only white	-0.007 (0.035)	-0.004 (0.038)	0.001 (0.031)	0.004 (0.018)
<i>Panel C: Effect of manufacturing exposure on relative employment</i>				
Baseline:				
Full period, 1900-1940	0.190*** (0.069)	0.169** (0.069)	0.205*** (0.076)	0.036 (0.058)
Mid period, 1900-1920	0.153*** (0.053)	0.121*** (0.031)	0.167** (0.070)	0.046 (0.062)
Alternative Exposure Measure:				
Electricity Intensity	0.150** (0.075)	0.134* (0.071)	0.153* (0.089)	0.020 (0.070)
Alternative Sample:				
Only U.S. natives	0.219*** (0.075)	0.195*** (0.072)	0.242*** (0.080)	0.047 (0.058)
Only white	0.188** (0.073)	0.168** (0.072)	0.208*** (0.080)	0.040 (0.058)

Note. Estimates from (2) in a sample of 310 occupations for Panels A-B and 201 occupations for Panel C with positive employment in the initial and final years of the period for the U.S. Census sample of males aged 16-64yrs in the labor force. Estimates are weighted by the occupation's employment at t_0 . Dependent variable is the change in the outcome and period indicated in each panel. Estimates in column (1) based on a single group of all males aged 16-64yrs. Estimates in columns (2) and (3) based on a pooled regression (including group-period fixed effects) with two groups: Younger group includes workers aged 16-29yrs and Older group includes workers aged 30-64 yrs. Column (4) reports the difference between the coefficients for younger workers in column (3) and older workers in column (2). The occupation exposure in Panels A-B is the occupation's average ICT-intensity across industries, except in the third row where it is the occupation cognitive intensity and in the fourth row where it is the occupation college share, and in Panel C is the occupation's average manufacturing-intensity across industries, except in the third row where it is the occupation's average electricity-intensity across industries. Exposure measures are normalized to have a unit standard deviation. Estimates in the last two rows of each panel use occupation outcomes computed from the sub-sample of U.S. native and white males. Robust standard errors in parentheses clustered by occupation with *** denoting a p-value <0.01, ** a p-value <0.05, and * a p-value <0.10.

Decomposition of employment reallocation across cohorts. We now relate the different patterns of occupation reallocation for older and younger workers in the two episodes to the occupational choices of successive cohorts. We do so by decomposing the reallocation of cohorts towards highly exposed occupations in each decade of an episode into a common decade effect for all cohorts (ζ_t) and cohort-specific deviations from that common effect ($\epsilon_{c,t}$). In particular, we estimate the following regression:

$$\log s_{Hc,t}^{\text{Emp}} - \log s_{Hc,t-10}^{\text{Emp}} = \zeta_t + \epsilon_{c,t}, \quad (\text{A.1})$$

where the dependent variable is the ten-year log-change in the share of individuals of cohort c that are employed in occupations with the 25% highest values of exposure (i.e., the set of highly exposed occupations that we denote by \mathcal{O}_H).²⁷ We consider a sample that includes each of the five U.S. Decennial Census years in the corresponding episode and, in any given year, each of the 49 cohorts aged 16-64 years old.

When the patterns of reallocation towards highly exposed occupations are similar across cohorts, the cohort-period residuals $\epsilon_{c,t}$ have a lower dispersion (i.e., the standard deviation of $\epsilon_{c,t}$ is lower), and account for a smaller share of the variation in the dependent variable (i.e., $1 - R^2$ of the regression is lower). In fact, in the extreme case in which all cohorts were to exhibit exactly the same employment growth in highly exposed occupations, we would have $\epsilon_{c,t} = 0$ and thus $1 - R^2 = 0$.

In [Table A2](#), we report the statistics associated with the importance of the cohort-specific components in explaining the employment reallocation towards the highly exposed occupations in each of the two episodes. Panel A shows that, compared to the recent reallocation to ICT-intensive jobs, cohort-specific effects explain less of the reallocation towards manufacturing-intensive jobs in the early 1900s – that is, both $1 - R^2$ and the standard deviation of $\epsilon_{c,t}$ are lower in the earlier episode. In Panel B, we show that these results remain valid when we control for cohort-specific time trends in the estimation of the decomposition in [\(A.1\)](#). Thus, our findings are not driven by changes across cohorts in their life-cycle of transitions to highly exposed occupations.

Cross-occupation evidence from other countries. We now investigate whether other developed countries experienced patterns of relative employment growth in occupations more exposed to ICT that are similar to those observed for the U.S. in recent decades. We rely on data from Eurostat for European countries and IPUMS International for Canada. Because of the coarseness of the data that is available for several countries, we consider nine occupation groups (2-digit ISCO occupations) and define “Younger” workers as those aged 15-39 years (with “Old” workers defined as those aged 40-64 years). Because

²⁷Formally, $s_{Hc,t}^{\text{Emp}} \equiv \sum_{o \in \mathcal{O}_H} \text{Emp}_{oc,t} / \sum_o \text{Emp}_{oc,t}$, where $\text{Emp}_{oc,t}$ is the number of individuals of cohort c employed in occupation o at year t .

Table A2: Importance of Cohort-specific Effects in Reallocation towards Highly Exposed Occupations in the Two Episodes

	Reallocation towards occupations intensive in	
	ICT, 1980-2019 (1)	manufacturing, 1900-1940 (2)
Panel A: Importance of cohort-specific components		
$1 - R^2$	0.968	0.359
St. Dev of $\epsilon_{c,t}$	0.200	0.087
Panel B: Importance of cohort-specific components (w/ cohort linear trend controls)		
$1 - R^2$	0.620	0.082
St. Dev of $\epsilon_{c,t}$	0.186	0.050
Cohort-period obs.	156	156
Cohort obs.	69	69

Note. Statistics obtained from the estimation of (A.1) in Panel A, and (A.1) with cohort-specific linear time trends in Panel B. For each column, the set of high-exposure occupations, \mathcal{O}_H , contains the top 25 percent of occupations in terms of the indicated exposure measure. The occupation exposure in column (1) is the occupation's average ICT-intensity across industries and in column (2) is the occupation's average manufacturing-intensity across industries. Data from U.S. Census sample of males aged 16-64yrs in the labor force.

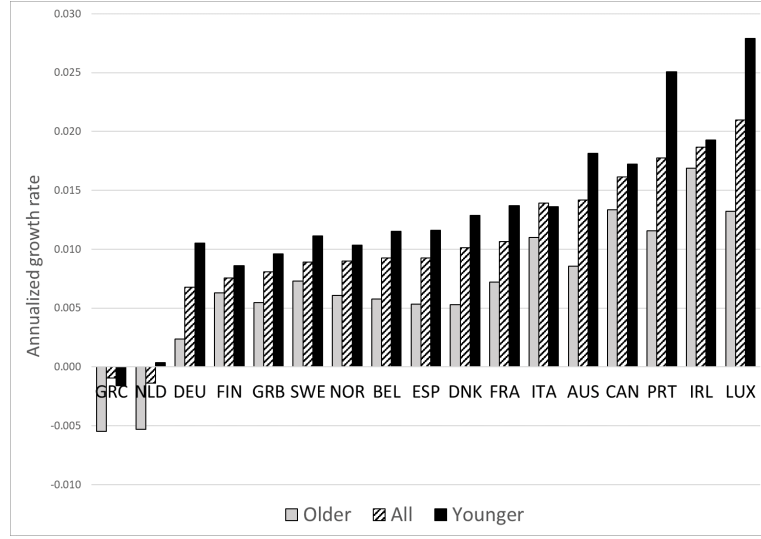
direct measures of ICT exposure are not available for occupations across all countries, we follow [Spitz-Oener \(2006\)](#) in constructing the cognitive intensity of occupations using the BERUFNET data, which is highly correlated with ICT exposure across occupations in the U.S.²⁸ The ICT-intensive occupations are the top three in this ranking: Managers, Professionals, Technicians and Associate Professionals.²⁹

[Figure A3](#) displays the recent trends of employment in ICT-intensive occupations for seventeen developed countries. The dashed bars indicate that relative employment in these occupations has been expanding in almost every country in our sample. This trend is a reflection of the occupation polarization process documented by [Goos et al. \(2009\)](#) and [Autor and Dorn \(2013\)](#). The figure also shows that, while older workers increased their relative employment in more exposed occupations in most countries, this increase was substantially stronger for younger generations. Across all countries, the annualized growth in the relative employment of younger workers was 73% higher than that of older workers. The young-old gap was higher whenever overall reallocation was higher: there is a correlation of 0.43 between the young-old gap in relative employment growth and that of all workers.

²⁸The BERUFNET dataset is based on the knowledge of experts about the skills required to perform tasks in each occupation. The occupation's cognitive intensity is the simple average of the time spent on analytical non-routine and interactive tasks in 2011-2013.

²⁹The other occupation groups are Elementary occupations, Plant and machine operators, Craft and related trades workers, Skilled agricultural and fishery workers, Service and sales workers, and Clerks.

Figure A3: Recent Trends in ICT-intensive Employment Growth in Developed Countries



Note. Annualized log-change in occupations more exposed to ICT in 1997-2017 for European countries, and in 2001-2011 for Canada. Sample of males in two age groups: “Younger” aged 15-39yrs and “Older” aged 40-64yrs. Occupations more exposed to ICT are Managers, Professionals, Technicians (out of the nine 2-digit ISCO occupation groups).

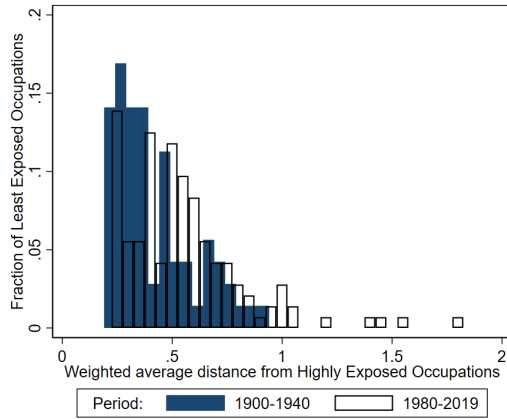
A.4 Additional Results: Differences in task content between occupations

Figure A4 assesses whether the patterns in Figure 2 are similar when we use alternative distance metrics, weighting, or exposure thresholds. Panel A defines the set of occupations exposed to the shock \mathcal{O}_H in each episode as the top 50% of all occupations in terms of exposure, with \mathcal{O}_L then defined as the remaining occupations. Panel B is the histogram of the bilateral distance, $D_{o,o'}$, between each occupation $o \in \mathcal{O}_L$ and $o' \in \mathcal{O}_H$. Panel C presents the histogram of task distances obtained with the alternative exposure measures in terms of electricity- and cognitive-intensity for the two episodes. Panel D presents results from considering the unweighted average task distance, with \mathcal{O}_H defined as the top 75% in terms of exposure.

In all cases, it is evident that the distribution of task distance for ICT exposure has more mass on higher values when compared to that implied by manufacturing exposure. For all specifications in Figure A4, we obtain a p-value of zero for the formal test of the first-order stochastic dominance of the distribution implied by ICT exposure in the second episode.

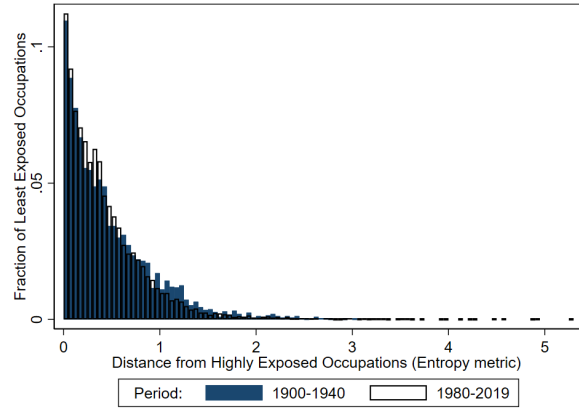
Figure A4: Task Content Distance Between Occupations with The Lowest and The Highest Levels of Exposure to the New Technologies in the Two Episodes – Alternative Specifications

A: \mathcal{O}_H defined for top 50% exposed occ.



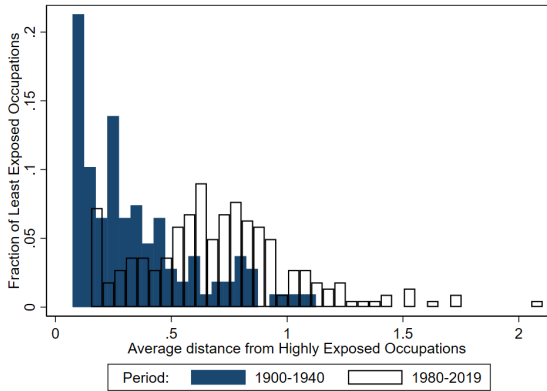
FOSD test: P-value 0.00, Difference 0.35***

B: Bilateral distance between all occ.



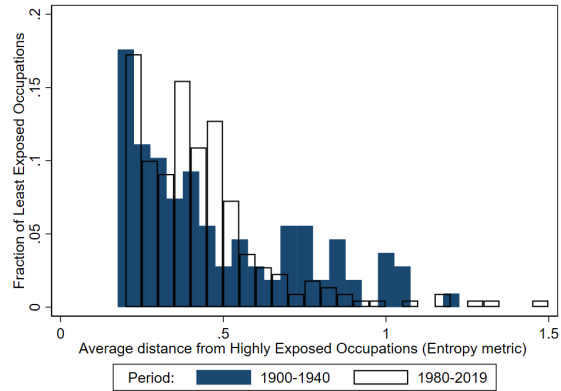
FOSD test: P-value 0.00, Difference 0.06***

C: Alternative exposure measures
 \mathcal{O}_H defined for top 25% exposed occ.



FOSD test: P-value 0.00, Difference 0.55***

D: Unweighted task distance
 \mathcal{O}_H defined for top 10% exposed occ.



FOSD test: P-value 0.01, Difference 0.19***

Note. The histograms plot the distributions of the average distance from high-exposure occupations in \mathcal{O}_H for each low-exposure occupation in \mathcal{O}_L , except in Panel B that plots the bilateral distance between each pair of occupations in \mathcal{O}_H and \mathcal{O}_L . Data from U.S. Census for employed males aged 16-64yrs. 1900-1940 uses the 1950 Census occupation definitions with positive employment in 1900 and 1940. 1980-2019 uses the 1990 occupation definition in Autor and Dorn (2009) with positive employment in 1980 and 2018. Distance between a pair of occupations is computed using an Entropy metric based on distance between the skill content of the pair of occupations using the Autor and Dorn (2013) measure of skills as discussed in Section 2. The sets of high and low exposure occupations, \mathcal{O}_H and \mathcal{O}_L , are respectively defined as the top 25 percent and the bottom 75 percent of occupations in terms of exposure in Panels B, C and D. In Panel A, the sets are defined as the top 50% and bottom 50% in terms of exposure. Panel C uses the alternative exposure measures to manufacturing in the early period and cognitive tasks in the latter period. Panel D uses an unweighted average distance for each \mathcal{O}_L with all occupations in \mathcal{O}_H , while Panels A and C report the employment weighted average distance. The reported test for first-order stochastic dominance is the Kolmogorov-Smirnov test.

A.5 Dynamic adjustment to ICT: Evidence from Germany

We now dig deeper into the differences across generations that we documented for the relative employment responses in occupations more exposed to ICT in the U.S. and other countries. Our analysis focuses on how the German labor market responded to one specific ICT innovation, broadband internet. This allows us to exploit two unique features. First, Germany has detailed information on technology usage on the job and workers' vocational training. Second, the rise in penetration of broadband internet in Germany was fast in the early 2000s, but there was quasi-random regional variation in broadband penetration (Falck et al., 2014).

This section is organized as follows. Section A.5.1 presents the data used in our analysis. We then turn in Section A.5.2 to the analog for Germany of our baseline specification of responses across occupations. Finally, Section A.5.3 reports our estimates of how early broadband adoption in a region affected relative employment in ICT-intensive occupations for different worker generations.

A.5.1 Data

Our primary source of labor market data in Germany is the LIAB Longitudinal Model. We follow Card et al. (2013) to construct a sample of employed males aged 15-64 that reside in West Germany. The raw data in the LIAB comes in the form of entire job histories of workers in the sample. Individual entries therefore contain worker information, as well as information on the start and end date of a job spell for that individual, the establishment location, and characteristics of the job spell. We transform this data into an annual panel dataset following the steps in Card et al. (2013). Specifically, we sequentially restrict our sample by selecting (i) males in West Germany, (ii) those aged 15-64 years at the time of the job spell, and (iii) the job-spell within a calendar year with maximum earnings. Finally, we impute each individual's district of employment using the district of the establishment of the individual's main job in each year. Since this information only exists after 1999, we assign the establishment's district in 1999 to all prior years.³⁰

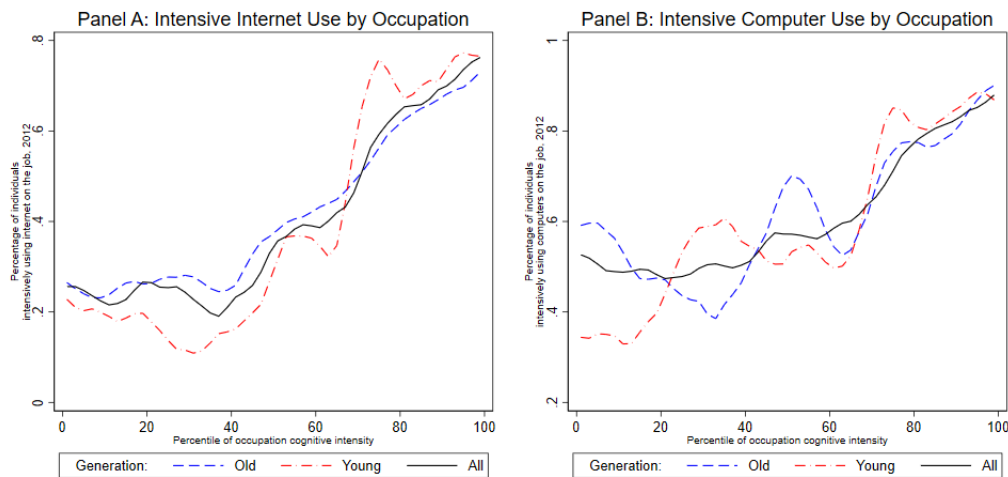
We use this dataset to construct yearly outcomes in 120 occupations for different groups of workers residing in the 323 districts of West Germany. We again consider two groups, younger and older workers, but we now define the younger group as those aged below forty. We choose a higher threshold than in our baseline for the U.S. because the age of labor market entry tends to be higher in Germany. We report that our estimates

³⁰While the years in our sample differ from those of Card et al. (2013), our panel well represents the data used in that paper. In results available upon request, we attest that the mean wage changes of job movers, classified by the mean log wages of coworkers in their old and new establishments, is similar in our data to the main findings in Card et al. (2013) (their Figure Vb).

are robust to alternative thresholds.

Lastly, we follow the analysis of [Spitz-Oener \(2006\)](#) for Germany to define ICT exposure as higher in occupations where individuals spend more time performing analytical non-routine and interactive tasks. Indeed, [Figure A5](#) graphically depicts how the use of recent ICT innovations varies with our exposure measure. It shows that usage of both internet and computers are strongly correlated with time spent on analytical non-routine and interactive tasks. Notice that internet and computer usages are not systematically different for younger and older workers conditional on their occupation's exposure.³¹ This suggests that the stronger reallocation of younger workers towards more exposed occupations that we documented in [Section 2.2](#) was not driven by differential technology usage by age given the skills required by an occupation.

Figure A5: Internet and Computer On-the-job Usage across Occupation



Note. Sample of 85 occupations in the 2012 Working Condition Survey. For each occupation, we compute the share of individuals reporting intensive internet and computer usage on their job. Figure reports the local polynomial smooth fit.

A.5.2 Exposure to new technologies and occupational outcomes

[Table A3](#) reports the estimation of equation (1) between 1995 and 2014 in our sample of 120 occupations in West Germany. Column (1) indicates that employment growth was 0.40% stronger for occupations in which exposure to ICT was one standard deviation higher. In line with the evidence for the U.S., columns (2) and (3) indicate that such an effect was stronger for younger workers than for older workers.³² Note that, compared to the U.S., estimates are larger in Germany, both for younger and older workers. This

³¹These results complement the finding in [Spitz-Oener \(2006\)](#) that there were small cross-cohort differences in the change of the task content of German occupations in the 1990s.

³²In results available upon request, we attest that that our estimates are similar when we either change the definition of younger generation, restrict the sample to native-born Germans, or control for exposure to trade and immigration shocks.

may be a consequence of the higher importance of high-tech manufacturing employment and on-the-job training in Germany.

We investigate next responses in skill investment by leveraging Germany’s large-scale vocational training program — an important component of the education of young workers, especially for those without a degree (Eckardt, 2019). Thus, the change in the number of trainees in an occupation is a proxy for changes in the skill investment decisions of young workers. To this end, column (4) reports the estimation of equation (1) where the dependent variable is the log-change in the number of trainees in occupation o . It shows that more exposed occupations experienced stronger growth in the number of trainees. This indicates that new workers increased their investment in the skills required by ICT-intensive occupations through vocational training.

Table A3: Change in Log Employment and ICT Exposure Across Occupations in Germany

Dependent variable:	Employment growth among			Trainee
	All	Younger	Older	Growth
	(1)	(2)	(3)	(4)
Exposure $_o$	0.398*** (0.060)	0.507*** (0.063)	0.233*** (0.061)	0.567*** (0.103)

Note. Estimates based on equation (1) between 1995 and 2014 in a sample of 120 occupations with positive employment among males aged 16-64 residing in West Germany. Robust standard errors in parentheses with *** denoting a p-value < 0.01

A.5.3 Dynamic adjustment of German local labor markets to broadband internet adoption

We now study how local labor markets in Germany adjusted to the arrival of one type of ICT innovation in the early 2000s, broadband internet. There are two main reasons to focus on this particular innovation in Germany. First, it resembles a permanent, one-time shock 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 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%. We follow Falck et al. (2014) to isolate exogenous spatial variation in adoption timing implied by the suitability of pre-existing local telephone networks for broadband internet transmission.

Empirical specification. For each year t , we consider the following specification:

$$\log \text{Emp}_{iog,t} - \log \text{Emp}_{iog,1999} = \sum_{c \in \{\text{young, old}\}} (\alpha_{c,t} + \beta_{c,t} \text{Exposure}_o) 1_{[g=c]} DSL_i + \zeta_{o,t} + \delta_{g,t} + X_{iog,t} \gamma_{g,t} + \epsilon_{iog,t} \quad (\text{A.2})$$

where $\text{Emp}_{iog,t}$ is the number of individuals of group g employed in occupation o and district i at year t , DSL_i is the standardized broadband internet penetration in district i in 2005, and Exposure_o is the occupation o 's exposure to ICT (as defined above). The specification includes group-year fixed effects that capture nationwide group-specific employment trends, as well as occupation-year fixed effects that absorb any shock affecting an occupation equally in all regions. We also include a control vector $X_{iog,t}$ to absorb potential confounding shocks associated the district's demographics.

The coefficient of interest is $\beta_{g,t}$: the differential impact of higher broadband penetration in 2005 in district i on the employment growth of occupations more exposed to ICT for workers of group g employed in that district. In all specifications, we consider 323 districts in West Germany and weigh observations by the district's employment.

Instrumental Variable. The consistent estimation of $\beta_{g,t}$ requires an exogenous source of variation on the adoption of broadband internet across districts. To this end, we follow Falck et al. (2014) by exploiting variation in the location of pre-existing main distribution frames (MDFs) of the telephone network. The initial roll-out of broadband internet in Germany in the early 2000s used the pre-existing telephone network. The transmission technology could not supply high-speed internet to areas more than 4200m away from an existing MDF. Thus, regions located close to MDFs were more likely to adopt broadband internet early. The use of MDF location as an instrument then requires that, conditional on controls, the determinants of MDF construction in the 1960s are orthogonal to the determinants of changes in labor market outcomes in the 2000s, except through their effect on broadband internet penetration in 2005. As argued by Falck et al. (2014), this is reasonable because the location of MDFs in the 1960s mainly depended on the availability of large empty sites, and did not take into account how attractive the region within a 4200m radius was for broadband internet suppliers in the 2000s (since MDFs provided telephone connections over much longer distances).

We then construct two instruments based on the district's average distance from existing MDFs, and interact them with group dummies and occupation exposure to estimate equation (A.2). The two instrumental variables are aggregates of the municipality-level variables in Falck et al. (2014) associated with the cost of locally supplying broadband internet through the existing telephone network in the early 2000s. The first variable ("MDF Density Measure") is a count of the number of municipalities in the district that did not have a MDF within the municipality, and whose population-weighted centroid

was further than the cut-off threshold of 4200m to the municipality's MDF. The second variable ("Alternative MDF Availability") counts the number of municipalities that satisfied the conditions in the first variable, but were further hampered by the lack of MDFs in neighboring municipalities that were closer than 4200m. [Figure A6](#) illustrates their spatial variation.

We now estimate the first-stage regression that relates the initial telephone network to the early expansion of broadband access across districts. Although the unit of observation in equation (A.2) is a district-occupation-generation triple, the exogenous variation in the instrument vector comes only from cross-district variation. Therefore, to provide a clear picture of the exogenous variation underlying the first-stage regression, we first examine the impact of the instrument vector Z_i on the district's share of population with broadband internet access in 2005, DSL_i .³³ That is, we begin by estimating the following linear regression:

$$DSL_i = Z_i\rho + X_i\gamma + \epsilon_i$$

where Z_i is the vector of instruments described above and X_i is the vector of district-level controls used in the estimation of (A.2).

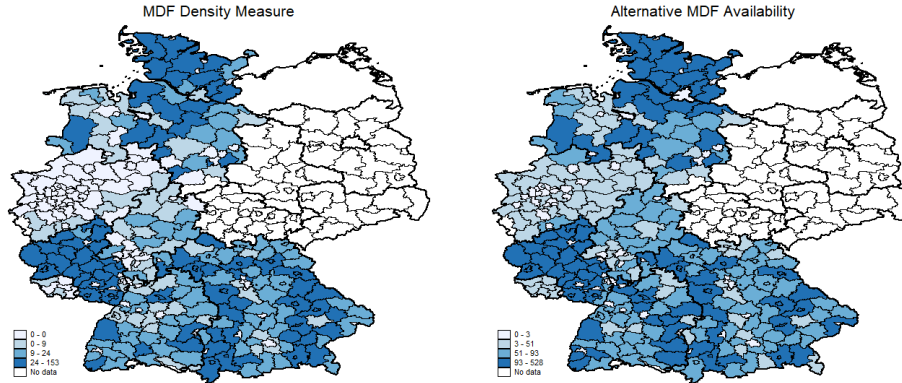
[Table A4](#) shows that districts with adverse initial conditions for internet adoption had a lower share of households with high-speed internet in 2005. Column (1) reports the first-stage estimates controlling for the baseline set of district-level controls. We can see that the F statistic of excluded variables remains high in the presence of these controls.

Since equation (A.2) includes broadband access interacted with occupation exposure and generation dummies, we present in [Table A5](#) tests for weak instruments in the presence of multiple endogenous variables. We obtain uniformly high first-stage Sanderson-Windmeijer F-statistics ([Sanderson and Windmeijer, 2016](#)) in all specifications, indicating that there are sufficiently many strong instruments for all endogenous variables.

Results. Panel A of [Figure A7](#) reports estimates of (A.2) for employed individuals in the younger and older groups. Before 2003, employment growth in more exposed occupations was similar in all regions for both worker groups. After 2006, higher broadband penetration caused stronger growth in the relative employment of younger workers in ICT-intensive occupations. Our estimates suggest that such an impact became stronger over time. In 2014, the point estimate indicates that a district with one-standard deviation higher broadband penetration in 2005 had 0.5 log-points higher employment of younger workers in the most exposed occupations (relative to the least exposed one). In contrast, we do not find any increase in the relative employment of older workers – if anything,

³³This is essentially the same as the expansion in broadband access between 2000 and 2005, since broadband internet was not available to households in Germany in 2000.

Figure A6: Spatial Variation in the Instrumental Variables



Note. Panel A illustrates the number of municipalities across districts in Germany that did not have access to an MDF within the 4200m radius ("MDF Density Measure"). Panel B illustrates the number of municipalities across districts that did not have their own MDF and did not have access to an alternative MDF in a neighboring district "Alternative MDF Availability".

Table A4: First-stage Regressions: Impact of Instrumental Variables on the Share of Households with Broadband Access in 2005

	(1)	(2)
MDF density measure	-0.020*** (0.005)	-0.018*** (0.005)
Alternative MDF availability	0.002 (0.001)	-0.001 (0.002)
Baseline controls	Yes	No
F statistic	26.49	43.06

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

Table A5: First-stage SW F-statistics for estimates of equation (A.2)

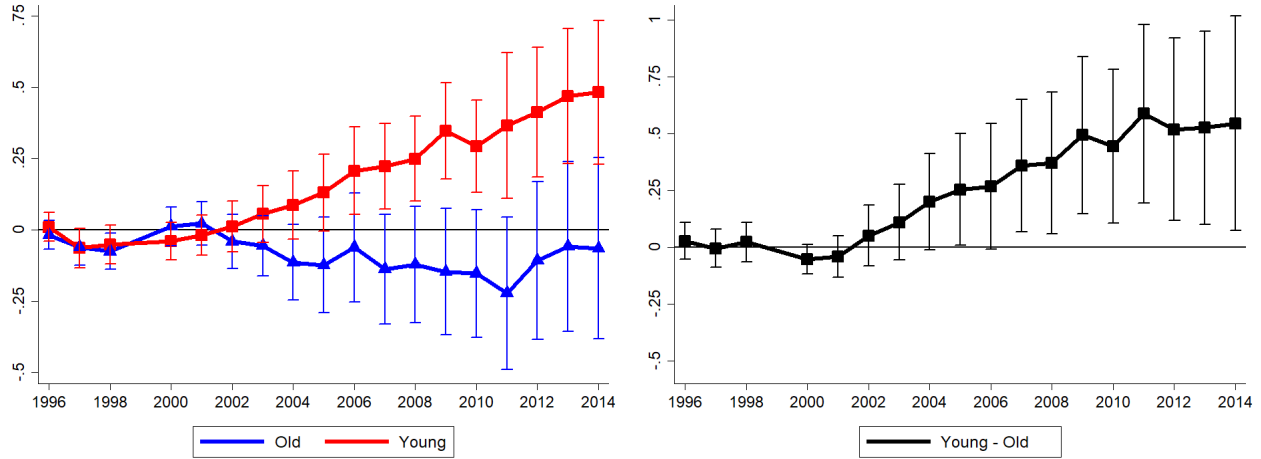
Instrumented Variable	1997	2007	2014
$Young_g \times DSL_i$	18.74	18.78	19.04
$Old_g \times DSL_i$	19.04	17.57	20.45
$Young_g \times DSL_i \times Exposure_o$	21.95	20.48	19.77
$Old_g \times DSL_i \times Exposure_o$	21.31	18.57	22.32

Note. Sample of 2 cohorts, 120 occupations and 323 districts. Table reports the Sanderson-Windmeijer F-statistic for each endogenous regressor when estimating equation (A.2).

it is negative. Panel B reports the difference between the impact of broadband access on the relative employment growth in more exposed occupations of younger and older generations, $\beta_{young,t} - \beta_{old,t}$. The between-generation gap in employment responses is

statistically significant at 10% for every year after 2006.

Figure A7: Impact of Early Broadband Expansion on Employment in Occupations More Exposed to ICT across German Districts and Generations



(a) Estimates of $\beta_{g,t}$ from equation (A.2)

(b) Estimate of $\beta_{young,t} - \beta_{old,t}$ from equation (A.2)

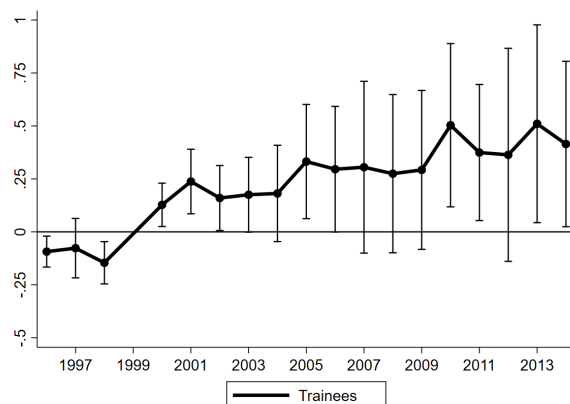
Note. Estimation of equation (A.2) in the sample of two generations (Younger and Older), 120 occupations and 323 districts in West Germany. The left panel reports $\beta_{g,t}$ for old and young generations, and the right panel reports $\beta_{young,t} - \beta_{old,t}$. All regressions are weighted by 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 associated 90% confidence interval implied by the standard error clustered at the district level.

We then investigate further what drives the stronger employment response of younger workers. We do so by estimating (A.2) when the outcome is the number of trainees in a single group with individuals of all ages in an occupation-district. Figure A8 shows that, prior to the reallocation of younger workers, there was an increase in the number of trainees in more exposed occupations during the roll-out period. Note that, in the early 2000s, the magnitude of the impact on the relative number of trainees is even larger than that on the relative employment of younger workers. But these two effects converge by the end of the sample period.

Robustness. We now turn to the robustness of our estimates of equation (A.2). We show that our results are robust to (i) dropping controls, (ii) controlling for district-generation-year fixed effects, (iii) using alternative definitions of the young generation, and (iv) using a sample without large establishments that might have received broadband internet early through private networks.

For our first set of robustness checks, we investigate in Table A6 how our set of baseline controls affects estimates. The three panels present estimates for the entire post-shock period of the sample (1999-2014, Panel A), the period during which broadband internet was rolled out across German regions (1999-2007, Panel B), and the period

Figure A8: Impact of Early Broadband Expansion on Number of Trainees in Occupations More Exposed to ICT across German Districts



Note. Estimation of equation (A.2) in the sample of 120 occupations and 323 districts in West Germany. The dependent variable is the number of employed trainees of all ages. 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.

before the shock (1996-1999, Panel C). Each panel includes the results of our baseline specification, as well as alternative specifications in which (i) we drop the pre-trend control, and (ii) we augment baseline controls with district-generation-year fixed effects.

Consider first the impact of the pretrend control in the second row of each panel. Results indicate that this control increases the magnitude and the precision of the estimates coefficients in the period of 1999-2007 and 1999-2014. However, it has the opposite impact on the pre-shock period of 1996-1999. In this pre-shock period, there are marginally significant negative responses without the pretrend control.

Turning to the specification including district-generation-year fixed effects, we can see that results are remarkably similar to our baseline estimates. This is reassuring as this specification includes a restrictive set of controls that absorb all potential confounding 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 exposure. That is, this control set captures any pre-existing variation that might have resulted in a district receiving broadband access early, including differential immigration into a district with early broadband expansion or differential aging or birth patterns in the district over time.

For our second set of robustness exercises, [Table A7](#) investigates the robustness of our estimates with respect to our sample selection choices. The second row of each panel shows that results are similar if we restrict the sample to only include workers born in Germany. This suggests that the inclusion of immigrants in our sample does not affect results.

We consider next several alternative definitions of the young generation based on (i) cohorts groups born after 1965 or 1970, and (ii) age groups aged below 35 or 45 in each year. For all definitions, relative employment responses for younger workers are positive and strongly significant, while that for older generations is non-significant and close to zero. Estimates in column (2) are stronger when we restrict the younger generation to cohorts born in more recent years or to have a lower age cutoff.

Finally, the last row of each panel reports estimates when we restrict the sample by excluding workers employed in establishments belonging to the top 25 percentile of the national distribution of establishment size. This exercise accounts for the likelihood that the largest establishments in Germany gained early broadband access through specialized private connections. In this case, we would expect adjustment in these establishments to have occurred earlier, biasing our results to zero. In line with this intuition, estimated coefficients are stronger than the baseline for all workers in column (1) and for the between-generation gap in column (4). This indicates that our instrument seems to generate variation in the roll-out of broadband internet that mostly affected the occupation composition of small establishments across German districts.

Table A6: Impact of Early Broadband Expansion on Employment in Occupations More Exposed to ICT across German Districts – Alternative Control Sets

Dependent variable: Control Set	Employment Growth			
	All (1)	Young (2)	Old (3)	Difference (4)
<i>Panel A: 1999-2014</i>				
Baseline	0.240*** (0.085)	0.482*** (0.154)	-0.065 (0.193)	0.546** (0.287)
No Pretrend Control	0.177** (0.087)	0.292*** (0.114)	-0.026 (0.189)	0.319 (0.222)
District-Year Effects	0.149** (0.067)	0.475*** (0.160)	-0.035 (0.203)	0.510* (0.302)
<i>Panel B: 1999-2007</i>				
Baseline	0.077* (0.043)	0.223*** (0.092)	-0.138 (0.116)	0.361** (0.177)
No Pretrend Control	0.015 (0.061)	0.137 (0.085)	-0.200 (0.127)	0.337 (0.149)
District-Year Effects	0.093 (0.060)	0.234** (0.098)	-0.134 (0.125)	0.368* (0.019)
<i>Panel C: 1996-1999</i>				
Baseline	-0.002 (0.026)	0.011 (0.030)	-0.019 (0.031)	0.029 (0.049)
No Pretrend Control	-0.109* (0.065)	-0.141* (0.077)	-0.074 (0.084)	-0.068 (0.061)
District-Year Effects	0.012 (0.032)	0.011 (0.032)	-0.022 (0.032)	0.034 (0.050)

Note. Sample of 2 cohorts, 120 occupations and 323 districts. Sample periods as defined in the table. Column (1) reports the estimated coefficient on the interaction between the occupation exposure and district broadband access in (A.2) for a single generation of working-age employed individuals. Columns (2)-(3) report the estimated coefficients on the interaction between the occupation exposure, generation dummies and district broadband access in (A.2) for the older and younger generations. Column (4) reports the difference between the coefficients in columns (3) and (2). All regressions are weighted by the district population size in 1999 and include a set of baseline district-level controls as well as occupation-year and generation-year fixed effects. Each row defines a separate robustness exercise. "District-Year Effects" are estimated with district-year fixed effects in column (1) and with district-year-generation fixed effects in columns (2)-(4). Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Impact of Early Broadband Expansion on Employment in Occupations More Exposed to ICT across German Districts – Sample selection

Dependent variable: Sample Definition	Employment Growth, 1999-2014			
	All (1)	Young (2)	Old (3)	Difference (4)
Baseline	0.240*** (0.085)	0.482*** (0.154)	-0.065 (0.193)	0.546** (0.287)
Native-born Males Only	0.223*** (0.078)	0.446*** (0.143)	0.074 (0.144)	0.372** (0.208)
Young: born after 1970		0.714*** (0.182)	-0.048 (0.242)	0.789** (0.371)
Young: born after 1965		0.612*** (0.157)	-0.171 (0.201)	0.783*** (0.303)
Young: Aged < 35 in each year		0.612*** (0.139)	0.059 (0.139)	0.553*** (0.203)
Young: Aged < 45 in each year		0.445*** (0.170)	0.159 (0.198)	0.286 (0.266)
Small Establishments Only	0.309*** (0.083)	0.466*** (0.140)	-0.128 (0.183)	0.594** (0.286)

Note. Sample of 2 cohorts, 120 occupations and 323 districts. Column (1) reports the estimated coefficient on the interaction between the occupation exposure and district broadband access in equation (A.2) for a single generation of working-age employed individuals. Columns (2)-(3) report the estimated coefficients on the interaction between the occupation exposure, generation dummies and district broadband access in equation (A.2) for the old and young generations. Column (4) reports the difference between the coefficients in columns (3) and (2). All regressions are weighted by the district population size in 1999 and include a set of baseline district-level controls, pretrend controls, occupation-year and generation-year fixed effects. Each row defines a separate sample selection exercise: (i) baseline sample restricted to only Germans ("Native-born"), (ii) different definitions of young workers based on year of birth or age cutoff in each year, and (iii) baseline sample restricted to workers employed in establishments below the 75th percentile of all establishment sizes ("Small Establishments Only"). Standard errors clustered at the district-level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix B Proofs

B.1 Proof of **Proposition 1** and **Proposition 2**

Part 1. Optimal assignment and labor supply of skill types. The optimal assignment in (10) is characterized by a threshold l_t corresponding to a skill type that is indifferent between working in the two technologies. The threshold satisfies

$$1 = \omega_t \sigma(l_t) = A_t^{\frac{\theta-1}{\theta}} \left(\frac{\int_{l_t}^1 \alpha(i) \sigma(i) s_t(i) di}{\sigma(l_t)^\theta \int_0^{l_t} \alpha(i) s_t(i) di} \right)^{-\frac{1}{\theta}}, \quad (\text{B.1})$$

where the last equality comes from market clearing. The right-hand side is strictly increasing in l_t , converges to infinity as $l_t \rightarrow 1$, and converges to zero as $l_t \rightarrow 0$. Then, existence and uniqueness of a solution follows from applying Bolzano's theorem.

The optimal labor supply of skill types in (11) is characterized by the FOC

$$\int_{\tau}^{\infty} e^{-\rho(t-\tau)} \frac{w_t(i)}{\int_0^1 w_t(x) \tilde{s}_\tau(x) dx} dt - (1 + \lambda_\tau) \left(\int_0^1 \tilde{s}_\tau(x)^{1+\frac{1}{\nu}} dx \right)^{\frac{1}{1+\frac{1}{\nu}}-1} \tilde{s}_\tau(i)^{\frac{1}{\nu}} = 0$$

$$\left(\int_0^1 \tilde{s}_\tau(x)^{1+\frac{1}{\nu}} dx \right)^{\frac{1}{1+\frac{1}{\nu}}} = 1.$$

Multiplying by $\tilde{s}_\tau(i)$ on both sides and integrating over $i \in [0, 1]$, we obtain $1 + \lambda_\tau = \frac{1}{\rho}$. Therefore,

$$\tilde{s}_\tau(i) = \left(\rho \int_{\tau}^{\infty} e^{-\rho(t-\tau)} \frac{w_t(i)}{\int_0^1 w_t(x) \tilde{s}_\tau(x) dx} dt \right)^\nu. \quad (\text{B.2})$$

Part 2. The q-theory system of ODEs. We begin by solving for the stationary distribution $s(i)$ implied by (12) and (B.2)

$$s(i) = \tilde{s}(i) = \frac{w(i)^\nu}{\left(\int_0^1 w(x)^{1+\nu} dx \right)^{\frac{\nu}{1+\nu}}} \quad (\text{B.3})$$

We then take a first order approximation around the stationary equilibrium of equa-

tions (12) and (B.1). We obtain

$$\frac{\partial \hat{s}_t(i)}{\partial t} = -\delta \hat{s}_t(i) + \delta \hat{\hat{s}}_t(i) \quad (\text{B.4})$$

$$\hat{l}_t = \frac{1}{\kappa} \frac{\eta}{\theta - 1} \hat{y}_t = \frac{1}{\kappa} \frac{\eta}{\theta} \hat{x}_t \quad (\text{B.5})$$

$$\hat{l}_t = \frac{1}{\kappa} \frac{\eta}{\eta + \theta} \left(\int_l^1 \hat{s}_t(i) \frac{\alpha(i)\sigma(i)s(i)}{\int_l^1 \alpha(i)\sigma(i)s(i)di} di - \int_0^l \hat{s}_t(i) \frac{\alpha(i)s(i)}{\int_0^l \alpha(i)s(i)di} di \right) \quad (\text{B.6})$$

where

$$\eta = \kappa \left(\frac{d \log(\sigma(x))}{d \log(x)} \Big|_{x=l} \right)^{-1}$$

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

Differentiating (B.6) with respect to time, we get that

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

Applying (B.4) to this expression and using the equation for wages (10), we obtain

$$\frac{\partial \hat{l}_t}{\partial t} = -\delta \hat{l}_t + \frac{1}{\kappa} \frac{\eta}{\eta + \theta} \delta \left(\int_l^1 \hat{s}_t(i) \frac{w(i)s(i)}{\int_l^1 w(i)s(i)di} di - \int_0^l \hat{s}_t(i) \frac{w(i)s(i)}{\int_0^l w(i)s(i)di} di \right). \quad (\text{B.7})$$

Log-linearizing (B.3), we obtain

$$\begin{aligned} \hat{\hat{s}}_\tau(i) &= \psi \rho \int_\tau^\infty e^{-\rho(t-\tau)} \left(\hat{w}_t(i) - \frac{1}{\int_0^1 w(x)s(x)dx} \int_0^1 w(x)s(x) (\hat{w}_t(x) + \hat{\hat{s}}_\tau(x)) dx \right) dt \\ &= \psi \rho \int_\tau^\infty e^{-\rho(t-\tau)} \left(\hat{w}_t(i) - \frac{1}{\int_0^1 w(x)s(x)dx} \int_0^1 w(x)s(x) \hat{w}_t(x) dx \right) dt, \end{aligned}$$

where the last equality follows from the fact that $\int_0^1 w(x)s(x) \hat{\hat{s}}_\tau(x) dx = 0$, since (B.3) holds and $\left(\int_0^1 \tilde{s}_\tau(x)^{1+\frac{1}{\nu}} dx \right)^{\frac{1}{1+\frac{1}{\nu}}} = 1$ for all τ .

Next, to obtain expressions for $\hat{w}_t(i)$, we guess and verify that l_t converges monotonically along the equilibrium path. We establish this starting from $\hat{l}_0 < 0$. We omit the analogous proof for $\hat{l}_0 > 0$. Whenever $\hat{l}_0 < 0$ and increases monotonically along the equilibrium path, we have that for all $s > t$, types $i < l_t$ are employed in technology L and types $i > l$ are employed in technology H . Also, for workers with $i \in (l_t, l)$, there

exist a $\tilde{\tau}(i)$ such that they work in H for all $\tau < t < \tau + \tilde{\tau}(i)$ and in L for all $t > \tau + \tilde{\tau}(i)$. This implies that

$$\hat{\omega}_t(i) = \left(\mathbb{I}_{i \geq l} + \mathbb{I}_{i \in (l_t, l)} \mathbb{I}_{t \in (\tau, \tau + \tilde{\tau}(i))} \right) \hat{\omega}_t.$$

Letting q_τ be the present discounted value of log-wages $\int_\tau^\infty e^{-\rho(t-\tau)} \log(\omega_t) dt$, we have that

$$\hat{q}_\tau = \int_\tau^\infty e^{-\rho(t-\tau)} \hat{\omega}_t dt,$$

and we obtain

$$\begin{aligned} \hat{s}_\tau(i) &= \nu\rho \left(\mathbb{I}_{i \geq l} - \frac{\int_l^1 w(x)s(x)dx}{\int_0^1 w(x)s(x)dx} \right) \hat{q}_\tau \\ &+ \nu\rho \left(\mathbb{I}_{i \in (l_t, l)} \left(\hat{q}_\tau - \hat{q}_{\tau + \tilde{\tau}(i)} \right) - \frac{1}{\int_0^1 w(x)s(x)dx} \int_{l_t}^l w(x)s(x) \left(\hat{q}_\tau - \hat{q}_{\tau + \tilde{\tau}(x)} \right) dx \right) \end{aligned}$$

Replacing this expression in (B.7), we obtain

$$\frac{\partial \hat{l}_t}{\partial t} = -\delta \hat{l}_t + \frac{1}{\kappa} \frac{\eta}{\eta + \theta} \delta \nu \rho \left(\hat{q}_\tau - \int_{l_t}^l \frac{w(x)s(x)}{\int_0^l w(i)s(i)di} \left(\hat{q}_\tau - \hat{q}_{\tau + \tilde{\tau}(x)} \right) dx \right).$$

Then, given our guess that l_t increases monotonically along the equilibrium path, from (B.1) we see that ω_t decreases monotonically along the equilibrium path. This implies that $\hat{q}_\tau > \hat{q}_{\tau + \tilde{\tau}(i)} > 0$ for all i and all τ . So, we can show that the term inside the integral is of second order:

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

Using (B.6) to replace \hat{l}_t , we obtain the law of motion for \hat{x}_t

$$\frac{\partial \hat{x}_t}{\partial t} = -\delta \hat{x}_t + \frac{\theta \nu \rho}{\eta + \theta} \delta \hat{q}_\tau \quad (\text{B.8})$$

Finally, we differentiate the definition of \hat{q}_t with respect to time

$$\rho \hat{q}_t = \omega_t + \frac{\partial \hat{q}_t}{\partial t}.$$

Part 3. The equivalence. We are now ready to show the equivalence of approximate equilibrium dynamics between the reduced-form and the micro-founded models. By

definition, (18) and (19) hold in both models. It then remains to be shown that (17) holds too.

Differentiating with respect to time and log-linearizing the relative labor supply equation in the reduced-form model, we obtain

$$\begin{aligned}
\frac{\partial \hat{x}_t}{\partial t} &= -\delta \hat{x}_t + \delta \hat{x}_{\tau,t} + e^{-\delta t} \frac{\partial \hat{x}_{0-t}}{\partial t} + \delta \int_0^t e^{-\delta(t-\tau)} \frac{\partial \hat{x}_{\tau,t}}{\partial t} d\tau \\
&= -\delta \hat{x}_t + \delta(\eta \hat{\omega}_t + \psi \hat{q}_t) + \eta \frac{\partial \hat{\omega}_t}{\partial t} \\
&= -\delta \hat{x}_t + \delta \left(-\frac{\eta}{\theta} \hat{x}_t + \psi \hat{q}_t \right) - \frac{\eta}{\theta} \frac{\partial \hat{x}_t}{\partial t} \\
&= -\delta \hat{x}_t + \frac{\theta \psi \delta}{\eta + \theta} \hat{q}_t
\end{aligned}$$

where the second and third equalities use (19). The last equation is (17) and is identical to (B.8) when $\psi = \nu\rho$; which completes the proof of equivalence.

Part 4. Saddle-path stability. We now show that the equilibrium is saddle-path stable, and verify that l_t increases monotonically along the equilibrium path.

We start by guessing that the policy functions are given by $\frac{\partial \hat{x}_t}{\partial t} = -\lambda \hat{x}_t$ and $\hat{q}_t = \zeta \hat{x}_t$. By replacing this guess into (17)–(19), we obtain the following system:

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

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

$$(\delta - \lambda)(\rho + \lambda) + \frac{\psi \delta}{\eta + \theta} = 0, \tag{B.9}$$

which yields the following solutions

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

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

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

completes the proof of the proposition.

B.2 Derivation of dynamic responses

Using the definitions x_t and q_t together with [Proposition 2](#), we have

$$\Delta \log(x_t) = (\theta - 1)\Delta \log(A) - \theta (\Delta \log(\omega) + \hat{\omega}_t) \quad (\text{B.10})$$

$$\Delta \log(q_t) = \frac{1}{\rho} \Delta \log(\omega) + \frac{1}{\rho + \lambda} \hat{\omega}_t \quad (\text{B.11})$$

Long-run. In this case, the stationary skill distribution is given by [\(B.3\)](#), so that the equilibrium threshold solves

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

A log-linear approximation around the stationary equilibrium gives

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

Thus,

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

From equation [\(B.1\)](#), $\Delta \log(\omega) = -\frac{\kappa}{\eta} \Delta \log(l)$ and, therefore,

$$\Delta \log(\omega) = \frac{1}{\theta + \frac{\psi}{\rho} + \eta} (\theta - 1) \Delta \log(A) \quad (\text{B.12})$$

Transition. We start by deriving the change in the skill distribution using [\(B.3\)](#): $\hat{s}_0(i) = \hat{s}_0(l)$ if $i < l$ and $\hat{s}_0(i) = \hat{s}_0(l) - \frac{\psi}{\rho} \Delta \log(\omega)$ if $i > l$. Along the transition, the change in the assignment threshold is determined by [\(B.1\)](#) given the change in the skill distribution:

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

Then,

$$\hat{\omega}_0 = -\frac{\kappa}{\eta} \hat{l}_0 = \frac{\psi}{\theta + \eta} \Delta \log(\omega) \quad (\text{B.13})$$

Dynamic responses. We now use the derivations above to show that

$$\Delta \log(x_t) = \left(\eta + \frac{\theta \frac{\psi}{\rho}}{\theta + \eta + \frac{\psi}{\rho}} (1 - e^{-\lambda t}) \right) \frac{\theta - 1}{\theta + \eta} \Delta \log(A) \quad (\text{B.14})$$

$$\Delta \log(\omega_t) = \left(1 - \frac{\frac{\psi}{\rho}}{\theta + \eta + \frac{\psi}{\rho}} (1 - e^{-\lambda t}) \right) \frac{\theta - 1}{\theta + \eta} \Delta \log(A) \quad (\text{B.15})$$

$$\Delta \log(x_{0-,t}) = \eta \Delta \log(\omega_t) - \frac{\psi}{\rho} \Delta \log(\omega) \quad (\text{B.16})$$

$$\Delta \log(x_{t,t}) = \eta \Delta \log(\omega_t) + \frac{\psi}{\rho} \left(1 + \frac{\lambda - \delta}{\delta} e^{-\lambda t} \right) \Delta \log(\omega) \quad (\text{B.17})$$

B.3 Comparative Statics with respect to η

Proposition A.1 (Comparative statics with respect to η) Assume that $\theta > 1$. Then,

1. *Short-run adjustment*

$$\frac{\partial \Delta \log(x_0)}{\partial \eta} > 0, \quad \frac{\partial |\Delta \log(\omega_0)|}{\partial \eta} < 0, \quad \frac{\partial \Delta \log(x_{0-,0})}{\partial \eta} > 0, \quad \frac{\partial \Delta \log(x_{0,0}/x_{0-,0})}{\partial \eta} < 0.$$

2. *Long-run adjustment*

$$\frac{\partial \Delta \log(x_\infty)}{\partial \eta} > 0, \quad \frac{\partial |\Delta \log(\omega_\infty)|}{\partial \eta} < 0, \quad \frac{\partial \Delta \log(x_{0-, \infty})}{\partial \eta} > 0, \quad \frac{\partial \Delta \log(x_{0,0}/x_{0-,0})}{\partial \eta} < 0.$$

3. *Rate of convergence*

$$\frac{\partial \lambda}{\partial \eta} < 0$$

4. *Cumulative impulse response*

$$\frac{\partial (\int_0^\infty |\hat{x}_t| dt)}{\partial \eta} < 0, \quad \frac{\partial (\int_0^\infty |\hat{\omega}_t| dt)}{\partial \eta} < 0.$$

Next, we prove each of the items of the proposition above.

1. Short-run adjustment

$$\begin{aligned}\Delta \log(x_0) &= \frac{\eta}{\theta + \eta}(\theta - 1)\Delta \log(A) \\ \Delta \log(\omega_0) &= \frac{\theta - 1}{\theta + \eta}\Delta \log(A) \\ \Delta \log(x_{0-,0}) &= \left(\frac{\eta}{\theta + \eta} - \frac{\frac{\psi}{\rho}}{\theta + \eta + \frac{\psi}{\rho}} \right) (\theta - 1)\Delta \log(A) \\ \Delta \log(x_{0,0}) &= \left(\frac{\eta}{\theta + \eta} + \frac{\lambda}{\delta} \frac{\frac{\psi}{\rho}}{\theta + \eta + \frac{\psi}{\rho}} \right) (\theta - 1)\Delta \log(A)\end{aligned}$$

The first and third lines immediately show that $\Delta \log(x_0), \Delta \log(x_{0-,0})$ increase with η , and the second line that $\Delta \log(\omega_0)$ decreases with it. Since λ is decreasing in η , the last line shows that $\Delta \log(x_{0,0}/x_{0-,0})$ decreases with η .

2. Long-run adjustment

$$\begin{aligned}\Delta \log(x_\infty) &= \frac{\eta + \frac{\psi}{\rho}}{\theta + \eta + \frac{\psi}{\rho}}(\theta - 1)\Delta \log(A) \\ \Delta \log(\omega_\infty) &= \frac{\theta - 1}{\theta + \eta + \frac{\psi}{\rho}}\Delta \log(A) \\ \Delta \log(x_{0-, \infty}) &= \left(\eta - \frac{\psi}{\rho} \right) \frac{\theta - 1}{\theta + \eta + \frac{\psi}{\rho}}\Delta \log(A) \\ \Delta \log(x_{0, \infty}) &= \left(\eta + \frac{\psi}{\rho} \right) \frac{\theta - 1}{\theta + \eta + \frac{\psi}{\rho}}\Delta \log(A)\end{aligned}$$

It is straightforward to see that $\Delta \log(x_\infty), \Delta \log(x_{0-, \infty})$ are increasing in η , while the opposite holds for $\Delta \log(\omega_\infty)$ and $\Delta \log(x_{0, \infty}/x_{0-, \infty})$.

3. Rate of convergence

From the implicit equation for λ in (B.9), it is straightforward to see that λ is decreasing in η .

4. Cumulative impulse response

$$\int_0^\infty |\hat{x}_t| dt = \theta \int_0^\infty |\hat{\omega}_t| dt \int_0^\infty |\hat{\omega}_t| dt = \frac{1}{\lambda} \frac{\frac{\psi}{\rho}}{\theta + \eta + \frac{\psi}{\rho}} \frac{\theta - 1}{\theta + \eta} \Delta \log(A)$$

To show that both are decreasing in η , we show that:

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

B.4 Proof of Proposition 4

We begin by showing the first order approximation to flow consumption utility $\log(C_{t,\tau})$. This equals the total flow utility $\log(C_{\tau,t}) - \rho L_\tau$ since $L_\tau = 1$.

$$\begin{aligned}
\log(C_{\tau,t}) &= \log\left(\int_0^1 \hat{c}_{\tau,t}(i) di\right) \\
&= \log\left(\int_0^1 w_t(i) \hat{s}_\tau(i) di\right) - \log(P_t) \\
&\approx \log(C) + \int_0^1 \frac{w(i)s(i)}{\int_0^1 w(x)s(x) dx} (\hat{w}_t(i) + \hat{s}_\tau(i)) di - \hat{P}_t \\
&= \log(C) + \int_0^1 \frac{w(i)s(i)}{\int_0^1 w(x)s(x) dx} \hat{w}_t(i) di - \hat{P}_t \\
&= \log(C) + \int_1^1 \frac{w(i)s(i)}{\int_0^1 w(x)s(x) dx} di \hat{w}_t - \hat{P}_t \\
&= \log(C) + \int_1^1 \frac{w(i)s(i)}{\int_0^1 w(x)s(x) dx} di \hat{w}_t - \frac{\omega x}{1+\omega x} \hat{w}_t \\
&= \log(C) + \frac{\omega x}{1+\omega x} \hat{w}_t - \frac{\omega x}{1+\omega x} \hat{w}_t \\
&= \log(C),
\end{aligned}$$

where $\Delta \log(C) = \frac{\omega x}{1+\omega x} \Delta \log(A)$. For the older generations, we have

$$\begin{aligned}
\log(C_{0-,t}) &\approx \log(C) + \int_0^1 \frac{w(i)s(i)}{\int_0^1 w(x)s(x)dx} (\hat{w}_t(i) + \hat{s}_0(i) - \hat{P}_t) di \\
&= \log(C) + \int_0^1 \frac{w(i)s(i)}{\int_0^1 w(x)s(x)dx} \hat{s}_0(i) di \\
&= \log(C) - \psi \int_0^1 \frac{w(i)s(i)}{\int_0^1 w(x)s(x)dx} \left(\mathbb{I}_{i>l} - \frac{\int_l^1 w(i)s(i)di}{\int_0^1 w(x)s(x)dx} \right) di \Delta \log(\omega) \\
&= \log(C).
\end{aligned}$$

The second order approximation is

$$\begin{aligned}
\log \left(\int_0^1 c_{\tau,t}(i) di \right) &\approx \log(C) + \int_0^1 \frac{c(i)}{C} \hat{c}_{\tau,t}(i) di - \frac{1}{2} \int_0^1 \left(\frac{c(i)}{C} \right)^2 \hat{c}_{\tau,t}(i)^2 di \\
\text{first order is zero} &\Rightarrow = \log(C) - \frac{1}{2} \int_0^1 \left(\frac{w(i)s(i)}{\int_0^1 w(x)s(x)dx} \right)^2 (\hat{w}_t(i) + \hat{s}_\tau(i) - \hat{P}_t)^2 di.
\end{aligned}$$

This implies

$$\begin{aligned}
\hat{C}_{\tau,t} &= -\frac{1}{2} \int_0^1 \left(\frac{w(i)s(i)}{\int_0^1 w(x)s(x)dx} \right)^2 \left(\mathbb{I}_{i \geq l} \hat{\omega}_t + \nu \left(\mathbb{I}_{i \geq l} - \frac{\omega x}{1+\omega x} \right) \rho \hat{q}_\tau - \frac{\omega x}{1+\omega x} \hat{\omega}_t \right)^2 di \\
\hat{C}_{\tau,t} &= -\frac{1}{2} \int_0^1 \left(\frac{w(i)s(i)}{\int_0^1 w(x)s(x)dx} \right)^2 \left(\mathbb{I}_{i>l} - \frac{\omega x}{1+\omega x} \right)^2 di (\hat{\omega}_t + \psi \hat{q}_\tau)^2 \\
&\propto -(\hat{\omega}_t + \psi \hat{q}_\tau)^2.
\end{aligned}$$

For older generations, we have

$$\begin{aligned}
\hat{C}_{0-,t} &\approx -\frac{1}{2} \int_0^1 \left(\frac{w(i)s(i)}{\int_0^1 w(x)s(x)dx} \right)^2 (\hat{w}_t(i) + \hat{s}_0(i) - \hat{P}_t)^2 di \\
&= -\frac{1}{2} \int_0^1 \left(\frac{w(i)s(i)}{\int_0^1 w(x)s(x)dx} \right)^2 \left(\mathbb{I}_{i>l} - \frac{\omega x}{1+\omega x} \right)^2 di \left(\hat{\omega}_t - \frac{\psi}{\rho} \Delta \log(\omega) \right)^2 \\
&\propto -\left(\hat{\omega}_t - \frac{\psi}{\rho} \Delta \log(\omega) \right)^2.
\end{aligned}$$

Then, aggregate welfare for all generations is

$$\begin{aligned}
\Delta W &= \rho(U_0 - U_{0-}) + \rho\delta \int_0^\infty e^{-\rho\tau} U_\tau d\tau \\
&= \Delta \log(C) + \rho \int_0^\infty e^{-\rho t} \hat{C}_{0-,t} dt + \rho\delta \int_0^\infty e^{-\rho\tau} \left(\log(C) + \int_\tau^\infty e^{-\rho(t-\tau)} \hat{C}_{\tau,t} dt \right) d\tau \\
&= \Delta \log(C)(1 + \delta) + \rho \left(\int_0^\infty e^{-\rho t} \hat{C}_{0-,t} dt + \delta \int_0^\infty e^{-\rho\tau} \int_\tau^\infty e^{-\rho(t-\tau)} \hat{C}_{\tau,t} dt d\tau \right).
\end{aligned}$$

Using the expressions above, we thus have that

$$\begin{aligned}
\Delta W - \Delta \log(C) \frac{\rho + \delta}{\rho} &\propto -(\hat{\omega}_0)^2 \rho \int_0^\infty e^{-\rho t} \left(e^{-\lambda t} - (\theta + \eta) \right)^2 dt \\
&\quad - (\hat{\omega}_0)^2 \rho\delta \int_0^\infty e^{-\rho\tau} e^{-2\lambda\tau} \left(\int_\tau^\infty e^{-\rho(t-\tau)} \left(e^{-\lambda(t-\tau)} + \frac{\psi}{\rho + \lambda} \right)^2 dt \right) d\tau \\
&= -(\hat{\omega}_0)^2 \left(\frac{\rho}{2\lambda + \rho} - 2 \frac{\rho}{\rho + \lambda} (\theta + \eta) + (\theta + \eta)^2 \right) \\
&\quad - (\hat{\omega}_0)^2 \frac{\rho\delta}{2\lambda + \rho} \left(\frac{1}{2\lambda + \rho} + \frac{2}{\rho + \lambda} \frac{\psi}{\rho + \lambda} + \frac{1}{\rho} \left(\frac{\psi}{\rho + \lambda} \right)^2 \right)
\end{aligned}$$

Finally, taking the limit $\rho \rightarrow 0$, we obtain

$$\begin{aligned}
\lim_{\rho \rightarrow 0} \frac{\partial \Delta W}{\partial \eta} &= \lim_{\rho \rightarrow 0} \frac{\partial \left(\Delta W - \Delta \log(C) \frac{\rho + \delta}{\rho} \right)}{\partial \eta} \\
&= -\frac{\partial}{\partial \eta} \left((\hat{\omega}_0)^2 \left((\theta + \eta)^2 + \frac{\delta}{2\lambda} \left(\frac{\psi}{\lambda} \right)^2 \right) \right) \\
&= -\frac{\partial}{\partial \eta} \left(((\theta - 1)\Delta \log(A))^2 \left(1 + \frac{\delta}{2\lambda} \left(\frac{\lambda - \delta}{\delta} \right)^2 \right) \right) \\
&= -((\theta - 1)\Delta \log(A))^2 \frac{1}{2\delta^2} \left(-\frac{1}{\lambda^2} (\lambda - \delta)^2 + \frac{1}{\lambda} 2(\lambda - \delta) \right) \frac{\partial \lambda}{\partial \eta} \\
&= -((\theta - 1)\Delta \log(A))^2 \frac{1}{2\delta^2} \frac{(\lambda - \delta)}{\lambda^2} (\lambda + \delta) \frac{\partial \lambda}{\partial \eta} > 0,
\end{aligned}$$

which completes the proof of the proposition.

Appendix C Additional Theoretical Results

C.1 Microfoundation of the Production Functions in (5)-(6)

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

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

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

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

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

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

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

To map this setting to the production functions in equations (5)-(6), we assume that high-tech production is more intensive in cognitive tasks than low-tech production, $\beta_H > \beta_L$. We also assume that types differ in terms of their skill bundle and, without loss of generality, impose that high- i types are relatively better in performing cognitive-intensive tasks.

1. High-tech technology H uses cognitive tasks more intensely than Low-tech technology L : $\beta_H > \beta_L$.
2. Define $\sigma(i) \equiv \left(\frac{a_C(i)}{a_{NC}(i)}\right)^{\beta_H - \beta_L}$ and $\alpha(i) \equiv a_C(i)^{\beta_L} a_{NC}(i)^{1-\beta_L}$. Assume that high- i types have higher cognitive-noncognitive task ratio: $\sigma(i)$ is increasing in i .

C.2 Extensions

This section discusses the extensions described in [Section 6.4](#).

C.2.1 Learning-from-others

We now assume that certain skills may be easier to supply because workers can "learn from others" when such types are already abundant in the economy. Formally, we assume that

$$L_\tau = \left(\int_0^1 \bar{s}_\tau(i)^{-\frac{1}{\nu}} \tilde{s}_\tau(i)^{\frac{1}{\nu}+1} di \right)^{\frac{1}{\frac{1}{\nu}+1}}$$

where $\bar{s}_\tau(i)$ is a geometric average of a fixed distribution $\bar{\epsilon}(i)$ and the current skill type distribution in the economy $s_\tau(i)$ at the time where generation τ is born,

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

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

In what follows, we reproduce the key steps that change in the proofs in [Appendix B.1](#). First, the optimal labor supply is

$$\tilde{s}_\tau(i) = \bar{s}_\tau(i) \left(\rho \int_\tau^\infty e^{-\rho(t-\tau)} \frac{w_t(i)}{\int_0^1 w_t(x) \tilde{s}_\tau(x) dx} dt \right)^\nu,$$

and the stationary distribution exist and is

$$s(i) = \frac{\epsilon(i)w(i)^{\frac{\nu}{1-\gamma}}}{\left(\int_0^1 \epsilon(x)w(x)^{1+\frac{\nu}{1-\gamma}} dx \right)^{\frac{\frac{\nu}{1-\gamma}}{1+\frac{\nu}{1-\gamma}}}}.$$

Following the same steps as in [Appendix B.1](#), we show that

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

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

$$(\delta(1-\gamma) - \lambda)(\rho + \lambda) + \frac{\psi\delta}{\eta + \theta} = 0.$$

Next, we reproduce the key steps that change in [Appendices B.2](#) and [B.3](#). First,

from the expression for the stationary distribution above, note that the long-run skill supply elasticity in the learning-from-others economy is $\frac{\psi}{1-\gamma}\psi$ as opposed to simply ψ . Likewise, in the expression for λ , the death rate is $\delta(1-\gamma)$ instead of just δ . In all, the learning-from-others economy behaves as if it had a higher ψ and lower δ .

C.2.2 Re-training of old workers

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

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

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

The first thing to note is that this does not change any of the transitional dynamics given the new initial skill distribution on impact. As such [Proposition 2](#) is unchanged. However, the initial conditions and the dynamic responses do change. Next, we reproduce the key steps that change in [Appendix B.2](#).

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

$$\begin{aligned}\hat{s}_0(i) &= \hat{s}_{0-}(i) + \beta(\hat{\tilde{s}}_0(i) - \hat{s}_{0-}(i)) \\ &= (1 - \beta)\left(\hat{s}_0(l) - \mathbb{I}_{i>l}\frac{\psi}{\rho}\Delta\log(\omega)\right) + \beta\left(\mathbb{I}_{i>l} - \int_l^1 s(i)di\right)\psi\hat{q}_0\end{aligned}$$

where the long-run change $\Delta\log(\omega)$ is the same as in the benchmark model. Following the same steps as in the benchmark proof, this then implies that

$$\begin{aligned}\frac{\theta + \eta}{\eta}\kappa\hat{l}_0 &= \int_l^1 \frac{\sigma(i)\alpha(i)s(i)}{\int_l^1 \sigma(i)\alpha(i)s(i)}\hat{s}_0(i)di - \int_0^l \frac{\alpha(i)s(i)}{\int_0^l \alpha(i)s(i)}\hat{s}_0(i)di \\ &= -(1 - \beta)\frac{\psi}{\rho}\Delta\log(\omega) + \beta\psi\hat{q}_0.\end{aligned}$$

Thus,

$$\begin{aligned}
\hat{\omega}_0 &= -\frac{\kappa}{\eta} \hat{l}_0 \\
&= \frac{1}{\theta + \eta} \left(\frac{\psi}{\rho} \Delta \log(\omega) - \beta \left(\frac{\psi}{\rho} \Delta \log(\omega) + \psi \hat{q}_0 \right) \right) \\
&= \frac{1}{\theta + \eta} \left(\frac{\psi}{\rho} \Delta \log(\omega) - \beta \left(\frac{\psi}{\rho} \Delta \log(\omega) + \frac{\psi}{\rho + \lambda} \hat{\omega}_0 \right) \right) \\
&= \frac{1 - \beta}{1 + \beta \frac{\psi}{\rho + \lambda} \frac{1}{\theta + \eta}} \frac{1}{\theta + \eta} \frac{\psi}{\rho} \Delta \log(\omega).
\end{aligned}$$

Finally, using the above together with the expression for $\Delta \log(\omega)$, we obtain:

$$\begin{aligned}
\Delta \log(x_t) &= \left(\eta + \frac{\theta \frac{\psi}{\rho}}{\theta + \eta + \frac{\psi}{\rho}} \left(1 - \frac{1 - \beta}{1 + \beta \frac{\lambda - \delta}{\delta}} e^{-\lambda t} \right) \right) \frac{\theta - 1}{\theta + \eta} \Delta \log(A) \\
\Delta \log(\omega_t) &= \left(1 - \frac{\frac{\psi}{\rho}}{\theta + \eta + \frac{\psi}{\rho}} \left(1 - \frac{1 - \beta}{1 + \beta \frac{\lambda - \delta}{\delta}} e^{-\lambda t} \right) \right) \frac{\theta - 1}{\theta + \eta} \Delta \log(A) \\
\Delta \log(x_{0-,t}) &= \eta \Delta \log(\omega_t) - \frac{\psi}{\rho} \Delta \log(\omega) \\
\Delta \log(x_{t,t}) &= \eta \Delta \log(\omega_t) + \frac{\psi}{\rho} \left(1 + \frac{\lambda - \delta}{\delta} \frac{1 - \beta}{1 + \beta \frac{\lambda - \delta}{\delta}} e^{-\lambda t} \right) \Delta \log(\omega)
\end{aligned}$$

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

C.2.3 Population growth

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

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

Then, we have that

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

The remaining elements in the model remain the same. Hence, the economy with population growth is identical to our benchmark economy except that the convergence rate λ is higher if, and only if, $\mu > \delta$ since it is now the positive solution to:

$$(\lambda - \mu)(\rho + \lambda) = \frac{\psi\mu}{\theta + \eta}.$$

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