Innovation in an Aging Economy*

Nils H. Lehr†

Boston University

October 2023

Abstract

This paper provides evidence that rapid workforce aging has contributed to slow productivity growth in the US over the last two decades, through its impact on innovation. I document that workforce aging in local labor markets leads to a reduction in R&D employment and fewer inventions using an instrumental variable strategy. Reductions in R&D employment are driven by within-age group changes, rather than a composition effect driven by age-specific R&D employment rates. This finding suggests that workforce aging impacts innovation through a demand channel, i.e., younger workers have a higher demand for inventions, rather than a supply channel operating, e.g., through the comparative advantage of young workers in innovation. Corroborating a strong demand channel of demographics, I also find that the workforce aging of international trading partners leads to a reduction in local innovation.

Keywords: Productivity slowdown, innovation, workforce aging, technology adoption

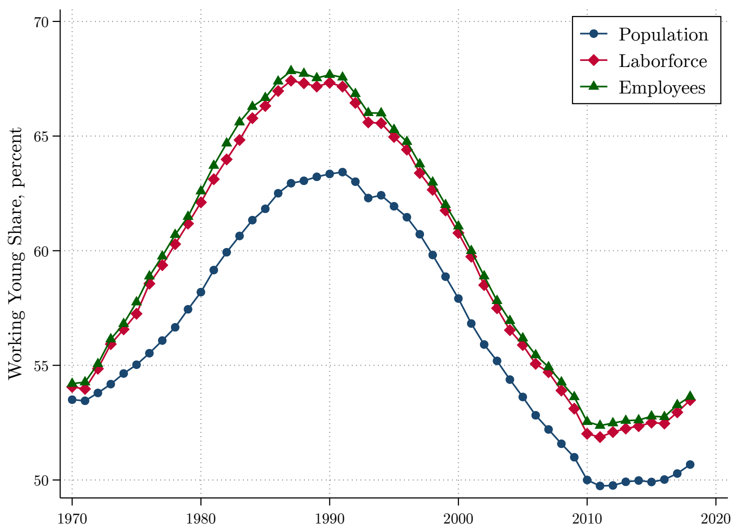
*I benefited from advice from many including Stephen Terry, Pascual Restrepo, David Lagakos, Daniele Paserman, Robert King, Tarek Hassan, and Susanto Basu as well as the useful comments and suggestions of participants in the Boston University Lunch Seminar, Green Line Macro Meeting, and Lecznar Memorial Lectures. Parts of this paper were previously circulated under the title "Aging, Technology Adoption, and Growth", which was first made publicly available in 2020.

†Email: nilslehr@bu.edu Mailing Address: Boston University, Dept. of Economics, 270 Bay State Rd., Room 515, Boston, MA 02215.

1 Introduction

The US workforce has aged significantly in recent decades, as have workforces in other developed nations. Figure 1 plots the share of people aged 25–44 among those aged 25–64, a measure I refer to as Working Young Share (WYS), for the US population, labor force, and employees. Between 1990 and 2010 the WYS for the US population has decreased from 63% to 50% with similar declines in absolute terms for the labor force and employees. The UN predicts the low WYS to persist in the medium to long-run implying that the workforce of the future is significantly older than the workforce of the past (Nations, 2019).

Figure 1: The US workforce aged rapidly since 1990



Note: This figure shows the WYS for the US population, labor force and employees based on the CPS ASEC samples. The WYS is defined as the share of age 25-44 subjects among those aged 25-64.

Concurrently, productivity growth has slowed down considerably over the last two decades (Syverson, 2017). Aghion et al. (2023) document that annual multifactor productivity growth in the US has declined from 0.79% in the 1988-95 period to 0.37% in the 2006-19 period, while Andrews et al. (2016) show that the productivity slowdown is a common phenomenon across developed nations. A growing literature attempts to explain the slowdown in productivity growth, however, few papers have explicitly linked it to demographics and workforce aging in particular (Teulings and Baldwin, 2014; Jones, 2019).

This paper presents evidence directly linking workforce aging to a reduction in innovation activity, which is often considered the main driver of medium- and long-run productivity growth (Romer, 1990). Following an instrumental variable approach, I find that an aging workforce in a local labor market leads to lower R&D investment, as measured by R&D employment, and fewer inventions, as measured by patents per capita. My estimation strategy proposes to address potential endogeneity concerns regarding changes in the local workforce age structure by leveraging historical birth-rates to create variation of workforce age composition that is arguably exogenous to subsequent changes in local labor market conditions.

Next, I investigate why workforce aging reduces innovation and find evidence suggesting that it is mostly due to a demand, i.e., market size, channel. I decompose changes in R&D employment rates into a within- and across-age group component, and find that the former accounts for 89% of the overall effect of workforce aging on R&D employment. Thus, workforce aging reduces R&D employment because fewer workers of a given age work in R&D and not because there are fewer workers in age groups with high R&D employment rates. This evidence is inconsistent with a supply-side interpretation in which workforce aging leads to less innovation because of young workers' comparative advantage in R&D. In such a world, we would expect the effect to be driven by the across age-group changes and might even see opposing within-group changes that equalize demand and supply for R&D workers. In contrast, the evidence is in line with a demand interpretation where young workers have more demand for new technologies, e.g., as they adopt them more frequently or faster, such that a larger share of young workers accelerates innovation due to a market size channel. In such a world, we would expect the overall effect to be dominated by within age-groups shifts in R&D employment rates, as I find empirically.

Finally, I provide additional, direct evidence for a demand-side channel by documenting that innovation activity declines significantly in response to workforce aging in export markets. Focusing on export markets mechanically shuts down any supply channel and, thus, allows for direct investigation of workforce aging as a shifter in the demand for innovation. Importantly, the documented link does not exists when investigating workforce aging in import markets, which confirms that export linkages do not capture trade connections in general, but sources of demand.

Together this evidence suggest that workforce aging reduces innovation due to a demand channel, which is qualitatively in line with an endogenous growth model with costly technology adoption and overlapping generations. In the model, older workers adopt fewer new technologies as they have less time remaining in the labor market to benefit from human capital investments. Thus, an economy with an older workforce has lower average technology adoption rates via a composition effect and, resultingly, lower demand for new technologies. This lack of demand then reduces investments in the creation of new technologies and, thus, productivity growth. Hence, economies with an aging workforce have slower economic growth.

I provide two pieces of evidence in line with this mechanism. Firstly, I show that older workers are slower to adopt new technologies using the rise of computers during the 1990s as an example. Secondly, I show that local labor markets with a rising share of young workers experience wage growth, but mostly so for young workers. This finding is in line with the model described above as the new technologies developed in light of a younger workforce raise the labor productivity for adopting workers, which tend to be young on average.

Literature. This paper contributes to three lines of research. Firstly, I contribute to the growing literature on the recent slowdown in US productivity growth by providing evidence in favor of workforce aging as a contributing factor. Syverson (2017) documents a significant slowdown in productivity growth since at least 2005, while Philippon and Gutiérrez (2017) documents a slow down in investment. A growing literature attributes this phenomenon to technology adoption lags (Brynjolfsson et al., 2019) or slower innovation (Jones, 2019; Bloom et al., 2020; Akcigit and Ates, 2021; Aghion et al., 2023; de Ridder, 2023; Liu et al., 2022). I complement the existing work by providing evidence for labor force aging as a contributing channel through its impact on innovation. This idea relates to the literature on firm dynamics and demographics, which argues that slower labor force growth, which is closely linked to aging, has contributed to declining firm dynamism (Hopenhayn et al., 2022; Karahan et al., 2022; Engbom, 2019; Peters and Walsh, 2022). I complement their analysis by providing evidence in favor of demand instead of supply-side factors as a core force shaping the impact of workforce aging on innovation and economic dynamism.

Secondly, the paper is closely related to the literature on the macroeconomic impact of aging, which has primarily focused on public finances and aggregate savings with three notable exceptions (Teulings and Baldwin, 2014; Eggertsson et al., 2019). Firstly, Feyrer (2007) and Maestas et al. (2023) provide evidence that labor force aging is associated with slower productivity growth at the state level.[1] My evidence suggests that this link can be partly explained by the impact of aging on innovation. Secondly, Aksoy et al. (2019) study population aging in a general equilibrium framework allowing for differential research productivity across age groups.[2] My estimates suggest that comparative advantage contributes little to the impact of workforce aging on innovation, which I find to be primarily driven by the demand-side. Finally, Acemoglu and Restrepo (2022) argue that workforce aging is a key driver of the current wave of automation and innovation therein. I complement their findings by providing evidence on the link between aging and overall innovation in the US, and by highlighting technology adoption as a potential underlying driving force.

Finally, I contribute to the literature connecting age to innovation and entrepreneurship by highlighting the demand-side implication of workforce aging on innovation. The literature documents that research and entrepreneurship productivity peaks around age 40–50, which suggests that workforce aging should increase entrepreneurship and R&D productivity (Jones, 2010; Jones and Weinberg, 2011; Ang and Madsen, 2015; Azoulay et al., 2020). In contrast, Derrien et al. (2023) find that local labor markets with a higher share of young workers record higher patenting rates. I extend this literature by, to the best of my knowledge, providing the first evidence on a causal link between workforce aging and investments in R&D. I also show that aging reduces innovation due to lower demand for new technology rather than comparative advantage across age groups, which has been the focus of this literature.

Organization. Section 2 discusses the potential effect of workforce aging on innovation. Section 3 introduces the data and empirical strategy followed by the results in Section 4. Section 5 concludes.

[1] Irmen and Litina (2022) and Aiyar et al. (2016) find similar patterns in the European and OECD context.

[2] Relatedly, Cai and Stoyanov (2016) and Gu and Stoyanov (2019) argue that young workers have more recent skills and investigate the implications for industry specialization.

2 Linking Aging to Innovation

There are at least two channels through which workforce aging might impact innovation and, thereby, productivity growth.[3] The first channel, which I refer to as a supply channel, posits that young workers are more apt to innovate.[4] Resultingly, an economy with a younger workforce has a larger relative supply of workers with a comparative advantage in innovation, which leads to a larger R&D sector as long as the economy is willing to substitute production and innovation at the margin.

The second channel, which I refer to as a demand channel, arises when young workers have a higher demand for inventions, i.e. the output of the innovation process and R&D sector. Resultingly, a younger workforce increases the market for innovation and potential profits thereof, which incentivizes firms to innovate more and leads to an expansion of the R&D sector. This channel is commonly referred to as market size effects in the growth literature (Jones, 1995).

One mechanism potentially giving young workers a higher demand for innovation is technology adoption. Young workers might adopt new technologies faster for three complementary reasons. First, they can reasonably expect to stay longer in the labor market and, thus, have more time to benefit from any investment in human capital through technology adoption. Secondly, young workers tend to have lower opportunity costs as are yet to have well developed expertise, e.g., in older technologies, or experience that pays them high wages.[5] Finally, older workers might be more focused on managerial tasks that rely less on technology and, therefore, be less likely to learn about and use new technologies (Acemoglu and Restrepo, 2022). For example, a senior software engineer might not need to learn about the latest developments in AI, because they rely on their junior colleagues for implementation. Independently of the particular channel, the literature finds that younger workers tend to adopt new technologies, or abandon outdated ones, at a faster rate (Friedberg, 2003; Weinberg, 2004; Horton and Tambe, 2020).

[3] Romer (1990) and Aghion and Howitt (1992) first argued that innovation is the main driver of long-run productivity growth. See, e.g., Kelly et al. (2021) for related evidence.

[4] The evidence on this assumption is mixed at best. For example, (Jones et al., 2014) and Azoulay et al. (2020) find that research and entrepreneurship productivity, respectively, peaks in middle ages.

[5] For example, Lagakos et al. (2018) document rising wages over the life-cycle and attribute the slope partly to human capital investments.

Online Appendix C proposes an endogenous growth model with overlapping generations and costly technology adoption that gives rise to such a demand channel. Older workers adopt fewer new technologies as they have less time in the labor market to take advantage of them. Resultingly, an older economy features less technology adoption, lower profits from innovation, and slower innovation-driven growth. Thus, workforce aging may slow down innovation and economic growth in theory, however, the extent to which it does in practice remains unclear.[6] In the following, I investigate this question empirically and find evidence for a strong demand channel.

3 Empirical Strategy

3.1 Data

My analysis links a range of data sources across time and space to investigate the impact of workforce aging on local innovation. My unit of analysis are 1990 US commuting zones (CZs) for decadal observations from 1980 to 2010 (Tolbert and Sizer, 1996). CZs are consistent geographic areas that are designed to capture a local labor market and are the standard geography considered in the literature on local labor markets. (Autor and Dorn, 2013; Autor et al., 2013) Unless otherwise noted, I map geographies to CZs using the crosswalk developed in Autor et al. (2013).

I construct employment- and population-based measures using the 1980, 1990, and 2002 decadal Censuses and the 2010-12 3-year ACS from IPUMS (Ruggles et al., 2020). I focus on full-time full-year (FTFY) workers, i.e., those reporting to have worked at least 40 weeks last year with at least 35 hours per week, for all employment-based measures following Acemoglu and Autor (2011). Occupations are measured using the consistent occupational codes developed in Autor and Dorn (2013).

I complement this data with information on local patenting from Berkes (2018), which is based on the USPTO's PatentsView. I map patents to CZs via the inventor's county of residence and split the credit for a patent equally among its inventors. I construct 5-year forward-citations via the patent citations file and define the technol-

[6] The model also highlights a potentially confounding channel: population growth. Faster population growth implies more potential customers in the future and, thus, a larger value of innovation today. Incidentally, faster population growth also implies a younger workforce. I discuss this challenge in the robustness section.

ogy class of a patent as its primary CPC sub-section. Forward-citations are citations received by a patent. As conventional in the literature, I record patents in their application year. (Kogan et al., 2017; Terry et al., 2023)

I obtain births by county from 1900 onward from historical censuses, the NBER Vitality Statistics, and the Surveillance, Epidemiology, and End Results (SEER) program, and map them to CZs via Eckert et al. (2018) and Autor and Dorn (2013).[7]

Finally, I obtain local employment by industry from the 2010-12 County Business Pattern and exports by industry in 2000 from the Census Foreign-Trade Statistics. I average values across years to safeguard against year-to-year fluctuations. I also obtain data on population size by age group by country from the UN database.

3.2 Measuring Local Innovation Activity

I create two measures of local innovation activity capturing innovation inputs and outputs, respectively. First, I proxy for local investments in innovation using the share of full-time full year (FTFY) workers in R&D occupations from the Census. I define the latter to be workers in natural sciences, engineering, social sciences, and computer science.[8] I consider this measure to be a reasonable proxy for local investment in innovation given that labor constitutes about 66.9% of total R&D cost according to the NSF's Business R&D and Innovation Survey and is, thus, an integral part of total R&D expenses. I refer to this variable as R&D employment.

Second, I measure innovation output using citation-weighted patents per 1,000 workers as in Terry et al. (2023). For this purpose, I first create a citation-based weight for each patent measuring the citations received relative to an average patent in the last 5 years and same technology class. I then attribute an equal share of each patent to its inventors and aggregate citation-weighted patents up to the CZ-year using their location and application year. Finally, I normalize citations-weighted patents for a CZ in a given year by the size of the local workforce to ensure my measure is comparable across CZs.

[7]In accordance with the terms of use of the Vital Statistics of the US as digitized by the NBER, I acknowledge indirect financial support from NIA grant P30-AG012810 through the NBER.

[8]Based on the consistent occupational codes developed in Autor and Dorn (2013), I classify four broad categories of occupations as R&D workers: natural scientists (codes 68-83), social scientists (166-169), computer scientists (64-65, 229-233), and engineers (44-59).

3.3 Approach

I investigate the effect of workforce aging on innovation in a simple first-difference specification for a local labor market g at time t following the growing literature on local labor markets (Autor and Dorn, 2013; Autor et al., 2013):

$$\Delta Y_{g,t} = \alpha_g + \gamma_t + \Delta X_{g,t} + \varepsilon_{g,t}, \tag{1}$$

where $Y_{g,t}$ and $X_{g,t}$ are measures of innovation and workforce aging, respectively, and Δ is the 10-year change in the variable. Estimating a difference specification safeguards against permanent differences across CZs driving my results and allows me to flexibly control for CZ-specific trends. I weigh observations by the CZ's initial working age population as in Autor et al. (2013).

Throughout, I focus on the WYS as my measure of workforce aging, which is defined as the ratio of the population age 25–44 to the population age 25–64 for a reference geography. The reference geography will be either the local CZ or the average over foreign nations to which a CZ is exposed via exports. I discuss the measure and its construction in greater detail together with the associated results.

4 Results

I present three results in this section. First, I document that local workforce aging reduces innovation using an instrumental variable strategy. While this result is an important first step, it is unclear whether it is driven by demand or supply factors. To shed light on the underlying drivers, I show that the link between local aging and R&D employment is driven by within age-group occupation changes rather than the mechanical effect of a shifting age distribution and argue that this result suggests demand factors as a driving force. Finally, I add to the evidence in favor of a strong demand-side link between aging and innovation by showing that local innovation also responds to workforce aging in export markets.

4.1 Local Workforce Aging and Innovation

The first step in my analysis is to investigate the link between workforce aging and innovation within a commuting zone. Let $\texttt{Pop}_{g,a,t}$ be the age a population in CZ g and year t, then the WYS is defined as

$$\texttt{WYS}_{g,t} = \frac{\sum_{a=25}^{44} \texttt{Pop}_{g,a,t}}{\sum_{a=25}^{64} \texttt{Pop}_{g,a,t}} \times 100. \tag{2}$$

Table A.1 reports summary statistics for changes in the local WYS. In line with Figure 1, the WYS has declined on average in my sample by 2.8 percentage points per decade. Notably, this decline has been relatively uniform across CZs. The unconditional standard deviation of changes in the WYS is around 6 percentage points, while it is only 1.7 percentage points once we take out year fixed effects.

A natural concern with estimating (1) in this context is reverse causality due to short-term shocks to local activity. For example, innovation might be positively correlated with other measures of labor market opportunities, which in turn could disproportionately attract young, more mobile workers. On the other hand, it could be the case that environments with a lot of innovation have high cost of housing, which might make them less attractive to young workers. In either case, we might expect a biased coefficient, however, the direction of the bias is unclear ex-ante.

To address this potential issue, I propose an instrumental variable strategy leveraging historical births following a growing literature on demographics in macroeconomics.[9] I define a hypothetical population $\widehat{\texttt{Pop}}_{g,a,t}$ as the share of total births in CZ g at time $t-a$ times the total US population of age a at time t:

$$\widehat{\texttt{Pop}}_{g,a,t} = \frac{\texttt{Births}_{g,t-a}}{\sum_g \texttt{Births}_{g,t-a}} \times \left(\sum_g \texttt{Pop}_{g,a,t} \right). \tag{3}$$

I then construct a WYS using this measure and equation (2) that relies upon variation in historical birth rates and, thus, isolates a long-term demographic component. Intuitively, the measure captures a WYS as if people never moved and were subject

[9] Similar identification strategies underlie empirical analyses in, e.g., Shimer (2001); Engbom (2019); Karahan et al. (2022); Derrien et al. (2023) and Acemoglu and Restrepo (2022). My instrument reaches further back in time by multiple decades as I construct births starting in 1900.

to identical mortality risks. By construction, the instrument should be unrelated to contemporaneous economic fluctuations that are not driven by the WYS itself.

Identification Assumption (Local WYS)**.** *Conditional on year and CZ fixed effects, changes in the local WYS instrument are only linked to innovation via the local WYS.*

Note that the instrument directly addresses the concern of differential worker mobility towards opportunities or properties of innovation environments that differentially affect young workers. The instrument does not address concerns that link birth rates to other characteristic of the CZ. For example, if there are differences in fertility rates across educational and ethnic groups, then this would be picked up by the instrument. I investigate this concern in the discussion section in a "bad control" exercise inspired by Angrist and Pischke (2009).

Table 1: Predicting the WYS from Births

	(1)	(2)	(3)
		$\Delta\texttt{WYS}$	
$\Delta\widehat{\texttt{WYS}}$	0.771***	0.168***	0.213***
	(0.022)	(0.032)	(0.035)
F statistic	552	42.5	37.2
Initial conditions	✓	✓	
CZ FE			✓
Year FE		✓	✓
Observations	2,166	2,166	2,166

Note: Inital conditions include the college, non-white, working young, and female share of the population as well as the metropolitan share and working age population size in 1980. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table 1 reports the first stage results. Column (1), which controls for a range of initial conditions, finds a strong relationship between the instrument and the WYS. A one percentage point increase in the hypothetical working young share is associated with an 0.78 percentage point increase in the actual WYS. This finding is not particularly surprising as the aggregate actual and hypothetical WYS coincide by construction. Once we control for year fixed effects in column (2), the coefficient drops significantly, but remains statistically and economically highly significant. Adding CZ fixed effects in column (3) does not materially affect the regression coefficient. The final first stage is strong with an F statistic around 31 and a 1 percent increase in the hypothetical WYS yields a 0.2 percentage points increase in the actual WYS.

Table 2: Local Workforce Aging and Innovation

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
A. Employment			**Δ R&D employment**			
ΔWYS	0.038***	0.029***	0.024	0.198***	0.025	0.217***
	(0.006)	(0.007)	(0.022)	(0.053)	(0.026)	(0.045)
B. Patenting			**Δ Citation-weighted patents**			
ΔWYS	0.037***	0.033***	0.019*	0.078**	0.023	0.092**
	(0.003)	(0.004)	(0.011)	(0.035)	(0.014)	(0.036)
F statistic		1,275		27.4		37.2
Initial conditions	✓	✓	✓	✓		
CZ FE					✓	✓
Year FE			✓	✓	✓	✓
Observations	2,166	2,166	2,166	2,166	2,166	2,166

Note: Inital conditions include the college, non-white, working young, and female share of the population as well as the metropolitan share and working age population size in 1980. First stage F statistics reported. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

With a strong instrument in hand, we can investigate the effect of local workforce aging on local R&D employment and patenting, which are reported in Table 2. Panel A reports the results for R&D employment. Controlling only for initial conditions in columns (1) and (2), we find a significant correlation between the WYS and R&D employment. A one percentage point increase in the WYS is associated with a 0.04 and 0.03 percentage point increase in the R&D employment share for the OLS and IV specification respectively. These results diverge once we add year and commuting zone fixed effects in column (3)-(4) and (5)-(6) respectively. While OLS results half in magnitude and become insignificant, IV results increase substantially. The IV coefficient from the full specification suggests that a one percentage point increase in the local WYS leads to a 0.19 percentage point increase in local R&D employment, equivalent to 0.25 standard deviations. The results for local patenting, as presented in Panel B, mirror those for local employment. In the full specification in column (6), a one percentage point increase in the WYS leads to 0.1 more citation-weighted patents per 1,000 workers or 0.2 standard deviations.

Overall, there are large differences between the OLS and IV results in Table 2, with the latter being much larger in absolute value. One possibility is that these differences are driven by reserve causality via a cost of living channel as suggested earlier. More generally, Shimer (2001) and Engbom (2019) also observe a large difference between OLS and IV using a similar identification strategy, but in a different context and with different outcome variables. Thus, this observation is known in the literature.

4.2 Innovation Demand as a Driving Force

As discussed in Section 2, a link between workforce aging and innovation can be explained by both demand and supply channels. To shed light on the underlying driving force, I propose a decomposition exercise that investigates whether movements R&D employment are driven by adjustments within or across age groups. I decompose changes in R&D employment into three terms:

$$\Delta \mathrm{Y}_{g,t} = \sum_{a=25,..,64} \underbrace{\Delta \mathrm{Y}_{a,g,t} \cdot \mathrm{P}_{a,g,t-1}}_{within} + \underbrace{\mathrm{Y}_{a,g,t-1} \cdot \Delta \mathrm{P}_{a,g,t}}_{across} + \underbrace{\Delta \mathrm{Y}_{a,g,t} \cdot \Delta \mathrm{P}_{a,g,t}}_{interaction},$$

where $Y_{a,g,t}$ is the share of local R&D workers among age a workers and $P_{a,g,t}$ is the share workers at age a. The *within* term captures changes in R&D employment within age groups, holding constant the age composition. The *across* term captures changes in the age composition holding constant age-specific R&D employment rates. The final term captures the *interaction* of age composition and R&D employment rates.

If the link between aging and innovation is primarily driven by comparative advantage of young workers in innovation, then one might expect R&D employment shares conditional on age to be unaffected by a changing WYS. Overall changes are then driven by the *across* term. In fact, one might even find a negative *within* coefficient as adjustments across groups are partly offset by the within-group margin to clear the labor market (Card and Lemieux, 2001). In contrast, changes in demand for innovation should lead a uniform increase in R&D employment across age groups and, thus, be reflected in the *within* term. Thus, regressing the individual terms on the changes in the WYS can shed light on the underlying mechanisms.[10]

Table 3: Local Workforce Aging and Innovation — Decomposition Results

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
A. Within Age	**Δ R&D employment**					
Δ`WYS`	0.023***	0.014*	0.014	0.172***	0.013	0.192***
	(0.006)	(0.007)	(0.022)	(0.056)	(0.025)	(0.045)
B. Across Age	**Δ R&D employment**					
Δ`WYS`	0.017***	0.017***	0.017***	0.027**	0.019***	0.033***
	(0.001)	(0.001)	(0.004)	(0.010)	(0.005)	(0.010)
C. Interaction	**Δ R&D employment**					
Δ`WYS`	-0.003***	-0.002**	-0.006***	-0.002	-0.007**	-0.008*
	(0.001)	(0.001)	(0.002)	(0.005)	(0.003)	(0.005)
F statistic		1,275		27.4		37.2
Initial conditions	✓	✓	✓	✓		
CZ FE					✓	✓
Year FE			✓	✓	✓	✓
Observations	2,166	2,166	2,166	2,166	2,166	2,166

Note: Inital conditions include the college, non-white, working young, and female share of the population as well as the metropolitan share and working age population size in 1980. First stage F statistics reported. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table 3 reports the associated results with Panel A, B, and C focusing on the *within*, *across*, and *interaction* term, respectively. Two results emerge. First, most of the effect of changes in the WYS on R&D employment is driven by the within-age group component. In the full specification, the within, across, and interaction term account for 89%, 15%, and -4% of the total effect, respectively. Second, the gap

[10]The *interaction* term is less informative about the relative strength of demand- and supply-side, however, it does have some information on the degree to which workers of different age groups are substitutes for each other. In particular, one might expect a negative coefficient if R&D workers are imperfect substitutes across age groups as the rise in one group's prevalence in the population is partly offset by a reduction in their R&D employment rates to balance changes in the overall age structure of R&D workers.

between OLS and IV is driven almost exclusively by the within component, which is small in the OLS and large in IV for the full specification. The gaps between IV and OLS are relatively small for the across and interaction component.

Thus, we can conclude that the causal effect of local workforce aging on innovation is not driven by a mechanical link between age and employment in R&D, as suggested by a comparative advantage-based narrative. Instead, the estimates suggest that demand factors are the primary driver of the workforce age to innovation link.

4.3 International Demand

Finally, to provide further evidence in favor of a strong demand channel, I investigate the impact of workforce aging in export markets on local innovation. The demand for new technologies is likely to extend beyond the local labor market. Thus, we can use changes in the WYS in markets that are sales destinations for local firms, but unconnected to their R&D labor market, as another test of whether there is a demand channel of workforce aging.

In particular, I propose to use international demand via exports as a source of variation in the WYS potentially driving the demand for local innovation. Intuitively, if there is a demand channel linking workforce aging and innovation, then it should be the case that firms linked to export destinations with an aging workforce reduce their innovation activity and vice versa. While direct commuting zone export data is not available, one can construct a proxy for workforce aging of international demand via industry-level exports and local industry employment. Let $\texttt{WYS}_{c,t}$ be the WYS in country c and $\texttt{Emp}_{g,i}$ be the employment of industry i in CZ g as measured in the County Business Patterns, I then calculate the industry level exposure as

$$\texttt{WYS}_{i,t}^{Trade,Ind} = \sum_{c} \left(\frac{\texttt{Exports}_{c,i}}{\sum_j \texttt{Exports}_{c,j}} \right) \times \texttt{WYS}_{c,t} \tag{4}$$

and map the industry-level measure back to the CZ using employment weights:

$$\texttt{WYS}_{g,t}^{Trade} = \sum_{i} \left(\frac{\texttt{Emp}_{g,i}}{\sum_{j \neq i} \texttt{Emp}_{g,j}} \right) \times \texttt{WYS}_{i,t}^{Trade,Ind}. \tag{5}$$

Regressing this measure on local innovation activity can then be interpreted as a reduced form estimate of the effect of workforce aging of international demand on local innovation. Variation in the proxy is driven by differential exposure of US industries to other countries' workforce aging and of CZs to US industries. For example, changes therein are particularly negative in CZs specialized in industries exporting to Japan.

Identification Assumption (International demand)**.** *Changes in the WYS across countries and as mapped to CZs via industry export and the CZ employment composition only affect CZs due to changes in the foreign WYS itself.*

Importantly, to identify a demand channel, I need to rule out that changes in the international WYS reflect other shocks that have an independent effect on innovation. Given that the WYS is defined at the national level, it appears unlikely that demand shocks simultaneously drive demand and the WYS. Thus, while confounding demand shocks might be a larger concern at a local level, they are not at a national level. However, one might be concerned that there are other channels at play that simultaneously influence US innovation and the WYS. For example, young countries tend to be poorer and less developed, giving them an overall lower demand for US innovation. Note, however, that I am focusing on changes in the WYS, which ignores level differences. Finally, by construction, I rule out any dynamic effects of exports and employment composition by fixing them over time.

Table 4 confirms a strong link between workforce aging of international demand and innovation. According to the most conservative specification in column (3), a one percentage point increase in the WYS in export markets is linked to a 0.1 percentage point higher R&D employment rate and 0.11 more forward-citations per 1000 workers. Importantly, the results are relatively stable across specifications. As in the previous exercise, these results suggests that workforce aging is linked to innovation due to a strong demand channel. Commuting zones exposed to workforce aging appear to reduce their innovation activity as they face lower demand for their inventions.

Table 4: Workforce Aging for Trading Parters and Innovation

	(1)	(2)	(3)
A. Employment	**Δ R&D employment**		
$\Delta\texttt{WYS}^{Trade}$	0.087***	0.070**	0.100***
	(0.010)	(0.035)	(0.034)
B. Patenting	**Δ Citation-weighted patents**		
$\Delta\texttt{WYS}^{Trade}$	0.083***	0.092***	0.118***
	(0.005)	(0.011)	(0.015)
Initial conditions	✓	✓	
CZ FE			✓
Year FE		✓	✓
Observations	2,166	2,166	2,166

Note: Inital conditions include the college, non-white, working young, and female share of the population as well as the metropolitan share and working age population size in 1980. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

4.4 Technology Adoption and the Demand for Innovation

How does the documented demand channel arise? As discussed in Section 2, one potential source is declining technology adoption rates over the life cycle. Two additional pieces of evidence support this idea. Firstly, Appendix D documents that older workers are slow to adopt new technologies by showing that older workers adopted the computer in the workplace at lower rates during the 1990s (Friedberg, 2003). Furthermore, the magnitudes are meaningful: workers age 55 and older had an about 15 pp lower propensity to use computers compared to the age 25-29 group.

Secondly, the technology adoption channel suggests faster wage growth in response to a younger workforce due to faster innovation, however, these effects should be more pronounced for younger workers due to their high technology adoption rates. To investigate this prediction, I construct wages and employment at the CZ level for all, young, and old workers separately, and then explore the impact of changes in the WYS on wage growth using the same estimation equation as for R&D employment.

Table 5: Workforce Aging and Wages

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
			Δ Wages			
ΔWYS	0.089***	0.142***	0.118***	0.218***	0.087***	0.085*
	(0.021)	(0.049)	(0.021)	(0.049)	(0.018)	(0.044)
Workers	All	All	Young	Young	Old	Old
CZ FEs	2,166	2,166	2,166	2,166	2,166	2,166

Note: All regressions control for CZ and year fixed effects. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

As documented in Table 5, the local WYS is linked to higher wages in general, however, its effect is especially pronounced for young workers. On average, a one percentage point increase in the WYS leads to a 0.14 percent increase in wages, however, the effect size almost doubles to 0.21 percent for younger workers, while old workers' wages only increase by 0.09 percent. Appendix Table B.1 confirms the same pattern for the WYS of international demand.

The evidence, thus, suggests rising wages in response to a younger workforce and higher relative wages for young workers. These findings are in line with the adoption channel: All wages increase as innovation improves labor productivity, while higher adoption rates among the young also increase their relative wages. Conversely, this finding is at odds with a simple supply shock interpretation as a rise in the relative supply of young workers should lead to lower, rather than higher, relative wages.

4.5 Discussion and Robustness

I document that CZs invest less in R&D and produce fewer inventions when faced with an aging workforce of demand. This finding is an important insight given the rapid workforce aging observed in the US and other developed nations. Before concluding, I consider a range of potential concerns that deserve further discussion.

General equilibrium effects. Adding year fixed effects in Table 2 raises the estimated coefficients significantly for the local WYS. One interpretation of this finding is that the first column incorporates general equilibrium effects or other aggregate shocks, while specifications with year fixed effects do not. In practice, we observe large changes in the WYS across all CZs over time, while changes in innovation are not as pronounced. Specifications with year fixed effects abstract from these baseline macro facts and, thus, put the focus on local variation. There are two ways to think about these differences. On the one hand, innovation partly has not declined as much in my sample due to the 1990s' dot-com boom. If we interpret this event as a one-time exogenous shock to R&D productivity that by chance occurred while the US workforce was aging, then the more stringent coefficient might be more accurate. On the other hand, the more stringent coefficient captures a partial equilibrium response due to year fixed effects, which might reflect a reallocation across CZs towards those with high technology demand as induced by the WYS, while CZs with relative low WYS are abandoned. This interpretation implies that simple calculation using the most stringent coefficients times the overall decline in the aggregate WYS could lead to misleading results. Similar considerations apply in the case of crowding out or changes in relative prices at the aggregate level. In this case, the coefficients still suggest an important role for the WYS in shaping the allocation of R&D, however, estimating the macro impact requires a full model.

Bad control. One concern with the instrumental variable strategy is that the instrument might be reasonably constructed yet unluckily reflect other drivers that are associated with birth rates and have a separate effect on innovation. I investigate this concern in a bad control exercise, where I first estimate the first stage and then use the predicted values in combination with other covariates to investigate the robustness of my estimates. A stable coefficient on the WYS should strengthen our confidence in the estimate (Angrist and Pischke, 2009).

$$\Delta Y_{g,t} = \alpha_g + \gamma_t + \beta \mathtt{WYS}_{gt}^{Pred} + \delta \Delta X_{g,t} + \varepsilon_{g,t} \tag{6}$$

I consider three confounders: gender, ethnic, and educational composition. Appendix Table B.5 confirms that neither the share of women, non-whites, or workers with bachelor degree explain the relationship between the WYS and innovation. The coefficient remains virtually unchanged when adding all three variables.

Age-composition vs population growth. Another issue of interpretation is the disentanglement of age composition and population growth, which are mechanically linked.[11] I attempt to address this issue by controlling for the growth rate of the population within the young generation. For this purpose, I define

$$\texttt{Pop Gr}_{g,t} = \left(\left(\frac{\sum_{a=25}^{34} \texttt{Pop}_{g,a,t}}{\sum_{a=35}^{44} \texttt{Pop}_{g,a,t}} \right)^{1/10} - 1 \right) \times 100, \tag{7}$$

and I construct an instrument using the same approach as for the WYS.

For export demand, I use the same mapping as for the WYS:

$$\texttt{Pop Gr}_{g,t}^{Trade} = \sum_i \left(\frac{\texttt{Emp}_{g,i}}{\sum_{j \neq i} \texttt{Emp}_{g,j}} \right) \times \sum_c \left(\frac{\texttt{Exports}_{c,i}}{\sum_j \texttt{Exports}_{c,j}} \right) \times \texttt{Pop Gr}_{c,t}. \tag{8}$$

Appendix Table B.3 reports the results for local WYS and population growth. The instrument is strong with an F statistic above 20 in the full specification, while the F statistic is above 50 for the WYS. Population growth itself has a positive impact on innovation as measured by R&D employment and patenting as documented in column (4). Importantly, the coefficients on the WYS are stable and significant when including population growth in the specification. The effect of the local WYS on innovation thus appears to be distinct from pure population growth. I reach a similar conclusion for the WYS of export markets. Appendix Table B.4 shows that the coefficient on $\texttt{WYS}^{Trade}$ is stable when controlling for population growth.

[11] To see this, suppose that there are only two generations alive at each point in time: young and old. Old workers leave the economy at the end of each period, while fraction $1-p$ young workers survive and become old. The size of the young generation grows at rate n. It is straight-forward to verify in this context that the overall population grows at rate n as well, while the WYS is given by

$$\texttt{WYS} = \frac{1+n}{2+n-p}.$$

It follows immediately that population growth and WYS are mechanically linked in the long-run. Furthermore, one can show that short-run fluctuation in p also link to short-run fluctuations in population growth. Thus, it is difficult to distinguish effects of the WYS and population growth separately.

Import linkages. Table 4 finds that workforce aging of export partners is associated with less innovation, which I interpreted as evidence in favor of a demand channel of workforce aging. However, one might be concerned that export partners also might be import partner such that these results could reflect a supply effect or knowledge spillovers. To address this concern, Table B.2 reports the international trade results using imports instead of exports. The full specification has a precisely estimated zero for employment and a significant negative coefficient for patenting. Thus, the results are specific to exports, in line with the demand interpretation.

Additional robustness checks. I conduct a range of additional robustness checks and report them in Appendix B. Firstly, I verify that my results are not driven by geographically correlated shocks by adding year×state fixed effects to the full specification. Appendix Table B.7 confirms that results for both measures of the WYS go through in this very stringent specification. Secondly, I report my main results unweighted and weighted by the beginning of period population in Appendix Tables B.8 and B.9. The results are essentially unaffected for the local WYS, while they are smaller in magnitude, but still significant, in the unweighted specifications for export demand. Thus, the qualitative conclusion is unaffected by weighting. Thirdly, I confirm in Appendix Tables B.10 and B.11 that my results are insensitive to an alternative definition of R&D workers excluding social scientists. Finally, Appendix Table B.6 reports results for alternative patenting measures and confirms that my baseline choice is not driving the result. The only exception is raw patenting, which has a positive, but insignificant, coefficient for the local WYS.

5 Conclusion

Over the last two decades the US has experienced fast workforce aging together with a slowdown in productivity growth. The share of young workers in the economy declined from 63% in 1990 to 50% in 2010, while productivity growth declined from 0.79% for the 1988–95 period to 0.37% in the 2006–19 period. This paper presents evidence suggesting that workforce aging led to less innovation and, thereby, slower productivity growth through a demand, i.e., market size, channel.

Firstly, I estimate that local labor markets experiencing faster workforce aging invest less in R&D and produce fewer inventions as measured by R&D employment and patenting, respectively, using an instrumental variable strategy. Thus, workforce aging leads to less innovation and, presumably, thereby to slower productivity growth.

Secondly, I show that the effect of workforce aging on R&D employment is driven by occupational changes away from R&D within age groups. This finding is at odds with a supply channel interpretation in which younger workers have a comparative advantage in innovation, which suggests that innovation and aging are linked due to the mechanical effect of a changing age distribution holding constant the age-specific R&D employment rates. Instead, the evidence is in line with workforce aging reducing the demand for innovation and thereby pushing workers out of the R&D sector, independently of their age.

Finally, I provide further evidence in favor of an important demand channel by estimating a decline in local innovation activity in response to workforce aging in export destinations. Commuting zones exposed to workforce aging in their export markets reduce their innovation, while the same does not hold true when focusing on import destination. Together, these results confirm that workforce aging reduces innovation due to its effect on the demand for innovation rather than the relative supply of workers with a comparative advantage in innovation.

These findings are qualitatively in line with a growth model of costly technology adoption and overlapping generations. In the model, older workers have lower technology adoption rates due to their limited time remaining in the workforce. Consequently, there will be less demand for new technologies in an aging economy and, thus, less innovation. These results suggest that workforce aging may have been an important contributor to slower economic growth. However, this slowdown might have been efficient to the degree that older workers optimally adopt fewer technologies and firms take this effect into account when deciding on their R&D investments. In other words, and similar to the argument presented in Vollrath (2020), slower economic growth in response to workforce aging might be welfare maximizing.

References

Acemoglu, Daron and David Autor, *Skills, Tasks and Technologies: Implications for Employment and Earnings*, Vol. 4, Elsevier,

— **and Pascual Restrepo**, "Modeling Automation," *AEA Papers and Proceedings*, 2018, *108*, 48–53.

— **and** — , "Demographics and Automation," *The Review of Economic Studies*, 2022, *89*, 1–44.

Aghion, Philippe and Peter Howitt, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, *60*, 323–351.

— **, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li**, "A Theory of Falling Growth and Rising Rents," *Review of Economic Studies*, 2023, *forthcoming.*

Aiyar, Shekhar, Christian Ebeke, and Xiaobo Shao, "The Impact of Workforce Aging on European Productivity," *IMF Working Papers*, 2016.

Akcigit, Ufuk and Sina T. Ates, "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," *American Economic Journal: Macroeconomics*, 2021, *13*, 257–298.

Aksoy, Yunus, Henrique S Basso, Ron P Smith, and Tobias Grasl, "Demographic structure and macroeconomic trends," *American Economic Journal: Macroeconomics*, 2019, *11*, 193–222.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal, "The Best Versus the Rest: The Global Productivity Slowdown, Divergence Across Firms and the Role of Public Policy," *OECD Productivity Working Papers*, 2016, *November.*

Ang, James B and Jakob B Madsen, "Imitation versus innovation in an aging society: international evidence since 1870," *Journal of Population Economics*, 2015, *28*, 299–327.

Angrist, Joshua D and Jorn-Steffen Pischke, *Mostly Harmless Econometrics* 2009.

Aubert, Patrick, Eve Caroli, and Muriel Roger, "New Technologies , Organisation and Age : Firm-Level Evidence," *The Economic Journal*, 2006, *116*, F73–F93.

Autor, David H. and David Dorn, "The growth of low-skill service jobs and the polarization of the US Labor Market," *American Economic Review*, 2013, *103*, 1553–1597.

_ , _ , and Gordon H. Hanson, "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 2013, *103*, 2121–2168.

_ , Frank Levy, and Richard J. Murnane, "The skill content of recent technological change: An empirical exploration," *Quarterly Journal of Economics*, 11 2003, *118*, 1279–1333.

_ , Lawrence F. Katz, and Alan B. Krueger, "Computing inequality: Have computers changed the labor market?," *Quarterly Journal of Economics*, 1998, *113*, 1169–1213.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda, "Age and High-Growth Entrepreneurship," *American Economics Review: Insights*, 2020, *2*, 65–82.

Bartel, Ann P. and Nachum Sicherman, "Technological change and the skill acquisition of young workers," *Journal of Labor Economics*, 1998, *16*, 718–755.

Berkes, Enrico, "Comprehensive Universe of U.S. Patents (CUSP): Data and Facts," *Working Paper*, 2018, pp. 1–27.

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

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt, "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-level Evidence," *The Quarterly Journal of Economics*, 2002, *117*, 339376.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics," *The Economics of Artificial Intelligence: An Agenda*, 2019, pp. 23–57.

_ , Lorin M. Hitt, and Shinkyu Yang, "Intangible assets: Computers and organizational capital," *Brookings Papers on Economic Activity*, 2002, *198*, 137–198.

Cai, Jie and Andrey Stoyanov, "Population aging and comparative advantage," *Journal of International Economics*, 9 2016, *102*, 1–21.

Card, D. and T. Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 5 2001, *116*, 705–746.

de Ridder, Maarten, "Market Power and Innovation in the Intangible Economy," 2023.

Derrien, Francois, Ambrus Kecskes, and Phuong-Anh Nguyen, "Labor Force Demographics and Corporate Innovation," *The Review of Financial Studies*, 2023, *36*, 2797–2838.

Eckert, Fabian, Andrés Gvirtz, and Michael Peters, "A Consistent County-Level Crosswalk for US Spatial Data since 1790 *," 2018.

Eggertsson, Gauti B., Neil R. Mehrotra, and Jacob A. Robbins, "A model of secular stagnation: Theory and quantitative evaluation," *American Economic Journal: Macroeconomics*, 2019, *11*, 1–48.

Engbom, Niklas, "Firm and Worker Dynamics in an Aging Labor Market," *Working Papers*, 2019, pp. 1–57.

Feyrer, James, "Demographics and Productivity," *The Review of Economics and Statistics*, 2007, *89*, 100–109.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren, "Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]," 2020.

Friedberg, Leora, "The Impact of Technological Change on Older Workers: Evidence from Data on Computers," *Industrial and Labor Relations Review*, 2003, *56*, 511–529.

Gancia, Gino and Fabrizio Zilibotti, "Horizontal Innovation in the Theory of Growth and Development," *Handbook of Economic Growth*, 2005, *1*, 111–170.

Gordon, Robert J, *The Rise and Fall of American Growth*, Princeton University Press, 2016.

Gu, Ke and Andrey Stoyanov, "Skills, population aging, and the pattern of international trade," *Review of International Economics*, 5 2019, *27*, 499–519.

Hopenhayn, Hugo A., Julian Neira, and Rish Singhania, "From Population Growth To Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share," *Econometrica*, 2022, *90*, 1879–1914.

Horton, John J and Prasanna Tambe, "The Death of a Technical Skill," 2020.

Irmen, Andreas and Anastasia Litina, "Population Aging and Inventive Activity," *Macroeconomic Dynamics*, 7 2022, *26*, 1127–1161.

Jones, Benjamin F., "Age and great invention," *Review of Economics and Statistics*, 2010, *92*, 1–14.

_ **and Bruce A. Weinberg**, "Age dynamics in scientific creativity," *PNAS*, 2011, *108*, 18910–18914.

_ **, E. J. Reedy, and Bruce A. Weinberg**, "Age and scientific genius," *The Wiley Handbook of Genius*, 2014, pp. 422–450.

Jones, Charles I., "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 1995, *103*, 759–784.

— , "The End of Economic Growth ? Unintended Consequences of a Declining Population," 2019.

Karahan, Fatih, Benjamin Pugsley, and Aysegül Sahin, "Demographic Origins of the Startup Deficit," 2022.

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

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth," *The Quarterly Journal of Economics*, 5 2017, *132*, 665–712.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman, "Life cycle wage growth across countries," *Journal of Political Economy*, 2018, *126*, 797–849.

Liu, Ernest, Atif Mian, and Amir Sufi, "Low Interest Rates, Market Power, and Productivity Growth," *Econometrica*, 2022, *90*, 193–221.

Maestas, Nicole, Kathleen J. Mullen, and David Powell, "The Effect of Population Aging on Economic Growth, The Labor Force And Productivity," *Americal Economic Journal: Macroeconomics*, 2023, *15*, 306–332.

Meyer, Jenny, "Older Workers and the Adoption of New Technologies," *ZEW - Centre for European Economic Research Discussion Paper No. 07-050*, 2007.

Nations, United, "2019 Revision of the World Population Prospects," 2019.

Ouimet, Paige and Rebecca Zarutskie, "Who works for startups? The relation between firm age, employee age, and growth," *Journal of Financial Economics*, 2014, *112*, 386–407.

Peters, Michael and Conor Walsh, "Population Growth and Firm Dynamics," 2022.

Philippon, Thomas and Germán Gutiérrez, "Investmentless Growth: An Empirical Investigation," *Brookings Papers on Economic Activity*, 2017.

Romer, Paul M, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98.*

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "IPUMS USA: Version 10.0 [dataset]," 2020.

Schleife, Katrin, "Empirical analyses of the digital divide in Germany - Age-specific and regional aspects," *Diss. Technische Universität*, 2008.

Shimer, R., "The Impact of Young Workers on the Aggregate Labor Market," *The Quarterly Journal of Economics*, 8 2001, *116*, 969–1007.

Syverson, Chad, "Challenges to Mismeasurement Explanations for the US Productivity Slowdown," *Journal of Economic Perspectives*, 5 2017, *31*, 165–186.

Terry, Stephen J., Konrad B. Burchardi, Thomas Chaney, Lisa Tarquinio, and Tarek A. Hassan, "Immigration, Innovation, and Growth," 2023.

Teulings, Coen and Richard Baldwin, *Secular Stagnation: Facts, Causes and Cures*, CEPR Press, 2014.

Tolbert, Charles M. and Molly Sizer, "US Commuting Zones and Labor Market Areas: A 1990 Update," 1996.

Vollrath, Dietrich, *Fully Grown*, Chicago University Press, 2020.

Weinberg, Bruce A, "Experience and Technology Adoption," *IZA DP No. 1051*, 2004.

Appendix

A Data

A.1 Data construction

Patenting measures. I construct citations weights for a patent as follows the 5-year forward citations divided by the average 5-year forward citations of granted patents with applied for in $t-4$ to t with the same primary CPC subsection. I split a patent equally if it has multiple inventors and assign each part to the CZ of the inventor. The total citation-weighted patents for a CZ is then simply the sum over the weighs times the splitting factors for all patents applied for in a given year by inventors residing in the CZ. To guard against outliers years I take the average of this measure over the $t-1$ to $t+1$ horizon. I normalize this value by the working age population in thousands via the Census.

Missing values. I impute missing values as 0. This applies to results for patenting and international trade based results. Results are robust to instead dropping the respective values.

Instrument. I construct historical births at the county level from three separate sources. Firstly, for the 1901-1939 period I rely on historical full-count census for 1910, 1920, 1930, and 1940.(Ruggles et al., 2020) For each decade I impute annual births using the age 0-9 population. The number of imputed births in a county in 1925 is thus the population born in 1925 as recorded in the 1930 census. Note that this naturally does not account for mortality up to age 5, an issue that I will discuss when detailing the actual construction of the instrument. For the 1940-67 period, I obtain births by county directly from the Vital Statistics of the US as digitized by the NBER. Finally, from 1967 onward I use the age 0 population recorded in the SEER data as my measure of births.[12] I map historical county-level birth to modern

[12]The digitized Vital Statistics of the US and SEER data are available here and here via the NBER.

CZs using the crosswalks developed in Eckert et al. (2018) and Autor and Dorn (2013).

I complement the data on births with data on the actual population size across age groups for the US from the NBER SEER data. The hypothetical population of age a at time t in CZ g is then the number of births in time $t-a$ divided by the total imputed births for the cohort times the actual population size of age a for the US in the particular year:

$$\widehat{\texttt{Pop}}_{g,a,t} = \frac{\texttt{Births}_{g,t-a}}{\sum_g \texttt{Births}_{g,t-a}} \times \left(\sum_g \texttt{Pop}_{g,a,t} \right). \tag{A.1}$$

By construction, this instrument gets the aggregate evolution of population groups correct, but uses historical births to distribute them across space. I aggregate this hypothetical measure across age groups to obtain the hypothetical WYS using the same formula as for the actual WYS itself:

$$\widehat{\texttt{WYS}}_{g,t} = \frac{\sum_{a=25}^{44} \widehat{\texttt{Pop}}_{g,a,t}}{\sum_{a=25}^{64} \widehat{\texttt{Pop}}_{g,a,t}}. \tag{A.2}$$

A.2 Summary statistics

Table A.1: Summary statistics

Variable	**Mean**	**SD**	**Within-year SD**
Δ R&D emp. (%)	0.492	0.829	0.666
Δ Age-adjusted R&D emp. (%)	-1.153	3.329	1.457
Δ Citation-weighted patents	-0.018	0.528	0.387
Δ Unbiased Citation-weighted patents	-0.040	0.553	0.399
Δ Unadjusted Citation-weighted patents	-0.021	0.569	0.445
Δ Patents	0.043	0.252	0.199
Δ Innovators	0.135	0.374	0.291
$\Delta\texttt{WYS}$	-2.976	5.912	1.730
$\Delta\widehat{\texttt{WYS}}$	-3.821	6.802	3.262
$\Delta\texttt{WYS}^{Trade}$	-1.758	2.695	0.488

Note: R&D employment in percentage points. Patenting values are per 1,000 workers. Final column residualizes variable with respect to year before calculating standard deviation. CZ observations weighted by 1980 working age population. See text for variable description.

B Robustness

Table B.1: Workforce Aging and Wages — Trade

	(1)	(2)	(3)
		Δ Wages	
$\Delta\texttt{WYS}^{Trade}$	0.227***	0.275***	0.193***
	(0.047)	(0.047)	(0.043)
Workers	All	Young	Old
CZ FEs	2,166	2,166	2,166

Note: All regressions control for CZ and year fixed effects. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.2: Workforce Aging for Trading Parters and Innovation — Import Robustness

	(1)	(2)	(3)
A. Employment	**Δ R&D employment**		
$\Delta \texttt{WYS}^{Trade}$	0.079***	-0.056*	0.003
	(0.009)	(0.030)	(0.033)
B. Patenting	**Δ Citation-weighted patents**		
$\Delta \texttt{WYS}^{Trade}$	0.073***	-0.063***	-0.081***
	(0.006)	(0.019)	(0.024)
Initial conditions	✓	✓	
CZ FEs			✓
Year FEs		✓	✓
Observations	2,166	2,166	2,166

Note: Inital conditions include the college, non-white, working young, and female share of the population as well as the metropolitan share and working age population size in 1980. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.3: Workforce Aging vs Population Growth

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
A. Employment			**Δ R&D employment**			
ΔWYS	0.023	0.170***			0.008	0.166***
	(0.022)	(0.040)			(0.023)	(0.037)
ΔPop gr			0.142***	0.239***	0.140***	0.133**
			(0.026)	(0.073)	(0.027)	(0.061)
B. Patenting			**Δ Citation-weighted patents**			
ΔWYS	0.023	0.092**			0.018	0.089**
	(0.014)	(0.036)			(0.016)	(0.033)
ΔPop gr			0.051**	0.152*	0.046**	0.095
			(0.020)	(0.089)	(0.022)	(0.074)
F stat. Δ WYS		37.2				55.2
F stat. Δ Pop. gr.				40.1		23.6
Observations	2,166	2,166	2,166	2,166	2,166	2,166

Note: All regressions control for CZ and year fixed effects. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.4: Workforce Aging vs Population Growth — Trade

	(1)	(2)	(3)
A. Employment	**Δ R&D employment**		
$\Delta\texttt{WYS}^{Trade}$	0.100***		0.072**
	(0.034)		(0.034)
$\Delta\texttt{Pop gr}^{Trade}$		-0.384***	-0.370***
		(0.060)	(0.059)
B. Patenting	**Δ Citation-weighted patents**		
$\Delta\texttt{WYS}^{Trade}$	0.118***		0.111***
	(0.015)		(0.014)
$\Delta\texttt{Pop gr}^{Trade}$		-0.118***	-0.095**
		(0.041)	(0.040)
Observations	2,166	2,166	2,166

Note: All regressions control for CZ and year fixed effects. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.5: Workforce Aging, Employment, and Wages

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
A. Employment		**Δ R&D employment**		
$\texttt{WYS}^{Pred}$	0.170***	0.169***	0.177***	0.166***
	(0.038)	(0.037)	(0.034)	(0.036)
Δ Non-white		0.005	0.006	0.008
		(0.010)	(0.011)	(0.010)
Δ Female			-0.124***	-0.074**
			(0.033)	(0.031)
Δ College				0.155***
				(0.028)
B. Patenting		**Δ Citation-weighted patents**		
$\texttt{WYS}^{Pred}$	0.092***	0.091***	0.093***	0.090***
	(0.032)	(0.032)	(0.031)	(0.031)
Δ Non-white		0.003	0.003	0.004
		(0.006)	(0.006)	(0.006)
Δ Female			-0.028	-0.014
			(0.024)	(0.026)
Δ College				0.043**
				(0.016)
Observations	2,166	2,166	2,166	2,166

Note: All regressions control for CZ and year fixed effects. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.6: Workforce Aging and Innovation — Alternative patenting measures

	(1)	(2)	(3)	(4)	(5)
A. Local	**Δ Innovation measure**				
Δ$\texttt{WYS}$	0.092**	0.083**	0.098***	0.027	0.057**
	(0.036)	(0.040)	(0.035)	(0.018)	(0.027)
B. Exports	**Δ Innovation measure**				
Δ$\texttt{WYS}^{Trade}$	0.118***	0.121***	0.129***	0.027***	0.049***
	(0.015)	(0.019)	(0.017)	(0.006)	(0.008)
Innovation measure	Baseline	Unadjusted	Unbiased	Patents	Inventors
Observations	2,166	2,166	2,166	2,166	2,166

Note: Column (1) reports the baseline results for comparison, column (2) uses raw Citation-weighted patents as weights, column (3) only counts citation from non-involved innovators, and column (4) does not ajust for citations at all. Finally, column (5) uses the number of inventors active in a CZ instead of patents. All regressions control for CZ and year fixed effects. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.7: Workforce Aging and Innovation — Results with State × Year FEs

	(1)	(2)
A. Employment	**Δ R&D employment**	
$\Delta\texttt{WYS}$	0.396***	
	(0.065)	
$\Delta\texttt{WYS}^{Trade}$		0.101***
		(0.028)
B. Patenting	**Δ Citation-weighted patents**	
$\Delta\texttt{WYS}$	0.163***	
	(0.036)	
$\Delta\texttt{WYS}^{Trade}$		0.077***
		(0.018)
Type		
Observations	2,157	2,157

Note: All regressions control for CZ and state× year fixed effects. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.8: Workforce Aging and Innovation — Unweighted results

	(1)	(2)
A. Employment	**Δ R&D employment**	
$\Delta\texttt{WYS}$	0.163***	
	(0.035)	
$\Delta\texttt{WYS}^{Trade}$		0.073***
		(0.014)
B. Patenting	**Δ 5-year Citation-weighted patents**	
$\Delta\texttt{WYS}$	0.076***	
	(0.019)	
$\Delta\texttt{WYS}^{Trade}$		0.043***
		(0.008)
Observations	2,166	2,166

Note: All regressions control for CZ and year fixed effects. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.9: Workforce Aging and Innovation — Alternative weight

	(1)	(2)
A. Employment	**Δ R&D employment**	
$\Delta\texttt{WYS}$	0.160***	
	(0.027)	
$\Delta\texttt{WYS}^{Trade}$		0.098***
		(0.029)
B. Patenting	**Δ Citation-weighted patents**	
$\Delta\texttt{WYS}$	0.076***	
	(0.018)	
$\Delta\texttt{WYS}^{Trade}$		0.117***
		(0.020)
Observations	2,166	2,166

Note: All regressions control for CZ and year fixed effects. Regressions weighted by beginning of period working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.10: Local Workforce Aging and R&D Employment Excluding Social Scientists

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
	Δ R&D employment excl. social scientists					
ΔWYS	0.026***	0.020***	0.018	0.151***	0.019	0.162***
	(0.005)	(0.006)	(0.017)	(0.053)	(0.020)	(0.039)
F statistic		1,275		27.4		37.2
Initial conditions	✓	✓	✓	✓		
CZ FE					✓	✓
Year FE			✓	✓	✓	✓
Observations	2,166	2,166	2,166	2,166	2,166	2,166

Note: Inital conditions include the college, non-white, working young, and female share of the population as well as the metropolitan share and working age population size in 1980. First stage F statistics reported. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Table B.11: Workforce Aging for Trading Parters and R&D Employment Excluding Social Scientists

	(1)	(2)	(3)
	Δ R&D emp. excl. soc. scientists		
$\Delta\texttt{WYS}^{Trade}$	0.063***	0.058*	0.082***
	(0.009)	(0.032)	(0.030)
Initial conditions	✓	✓	
CZ FE			✓
Year FE		✓	✓
Observations	2,166	2,166	2,166

Note: Inital conditions include the college, non-white, working young, and female share of the population as well as the metropolitan share and working age population size in 1980. CZ observations weighted by 1980 working age population. Standard errors clustered at the state level. See text for variable description.

Standard errors in parentheses. Significance levels: * 10% , ** 5%, *** 1%.

Online Appendix

Not for publication

C Aging, Technology Adoption, and Growth

This section develops an endogenous growth model that features a direct link between workforce aging and innovation via a demand channel. This feature allows it qualitatively to capture some of the patterns documents in Section 4. The model builds on the standard expanding varieties growth model as in Romer (1990) and extends it in two directions.[13] Firstly, I introduce demographics using a standard overlapping generations structure, and, secondly, technology adoption is made an explicit choice on part of workers.

C.1 Environment

Time is discrete and indexed by t. The economy features four types of agents. Households work, learn about technologies, and face a standard savings-consumption choice. The final goods sector in turn hires workers and buys equipment at competitive prices to produce the final good. Equipment is produced by specialized monopolists using the final good. Finally, new equipment varieties, which I will refer to as new technologies, are produced by an innovation sector, which borrows from households and repays them using profits generated by the associated equipment manufacturers. The final good is chosen as the numeraire.

I will denote the set of available technologies and new inventions as A_t and a_t respectively. The stock of technologies evolves cumulatively by adding new inventions:

$$A_t = a_t + A_{t-1}. \tag{C.1}$$

Households. The representative household maximizes

[13] See Gancia and Zilibotti (2005) for an introduction to expanding variety growth models.

$$\sum_{s=0}^{\infty} \beta^s (1+n)^s \ln(c_{t+s}), \tag{C.2}$$

where β is the time discount factor, n is the population growth rate, and c_t is per capita consumption.[14]

The household derives income from interest r_t on savings b_t and wages w_t, and spends it on savings, consumption, and technology adoption h_t. Technology adoption is linked to labor income and will be discussed in detail below. I focus on per capita values throughout to simplify the exposition. The budget constraint is given by

$$(1+n)b_{t+1} = (1+r_t)b_t + w_t - h_t - c_t. \tag{C.3}$$

Savings are restricted to be non-negative, $b_{t+1} \geq 0$.

The household is composed of two generations, young and old. The old generation exits the economy at the end of each period. It is replaced by the current young generation, whereof a share $1-p$ survives across periods. The young generation is replaced by a new young generation whose size grows at rate n. The setup gives rise to a constant share of young workers in the economy, denoted by s_y:

$$s_y = \frac{1+n}{2+n-p} \tag{C.4}$$

The analysis below focuses on comparative statics with respect to the population growth rate n and abstracts from transition dynamics induced by time-varying birth rates. Comparative statics for n are the appropriate analysis when considering the US. As discussed in Engbom (2019) and Karahan et al. (2022), the demographic patterns in Figure 1 are primarily driven by declining fertility rates.

Technology adoption is modeled as a costly, one-off investment on part of the household. Each period the representative household is confronted with the set of available technologies and decides for each worker which additional technologies to adopt. There is no forgetting, so a worker will be able to use a skill for the rest of her life once learned. Furthermore, workers can supply one unit of labor for all

[14] Log utility is chosen to keep the exposition simple and can be replaced by a CRRA utility function without changing the main results. I will throughout assume $\beta(1+n) < 1$ to ensure effective discounting on part of the household.

technologies in their skill set, so a larger skill set translates into a larger effective labor supply.

For technology $a \in A_t$ let $\ell_t(a)$ be the share of workers in the economy that have adopted the technology and $\ell_{gt}(a)$ be the share of workers of age group g that have adopted the technology. The former is then simply a weighted average of the latter:

$$\ell_t(a) = s_y \ell_{yt}(a) + (1 - s_y)\ell_{ot}(a). \tag{C.5}$$

Labor supply earns technology-specific wage $W_t(a)$. Per capita labor earnings are given by

$$w_t = s_y \int_{A_t} \ell_{yt}(a) W_t(a) da + (1 - s_y) \int_{A_t} \ell_{ot}(a) W_t(a) da \tag{C.6}$$

Knowledge does not come for free, however. All technologies are subject to per worker learning costs, which are i.i.d. distributed across technologies and workers, and constant over time for a particular technology-worker combination. I will denote the distribution by $F(n)$, where n is the cost of adopting a particular technology in terms of final goods. Workers do not differ in their inherent learning ability. Thus, I abstract from any considerations of reduced learning ability over the life-cycle or similar mechanisms.[15]

From the perspective of the household, workers in a given cohort look identical except for the technology adoption costs. Furthermore, I will show below that in equilibrium we will have $W_t(a) = W_t$ such that technologies will look identical from the perspective of a worker apart from their adoption costs. This facilitates the analysis greatly, as we can focus on adoption costs only.

Cohorts enter the economy with a blank slate and, thus, available technologies are indistinguishable to them apart from their adoption costs. We can thus think of the household's optimization problem as choosing a threshold type n_{yt} such that young workers adopt all technologies with cost type $n \leq n_{yt}$. The total adoption costs per young worker h_{yt} and effective labor supply for a technology $\ell_{yt}(a)$ are thus given by

[15] It is straight-forward to incorporate them and they amplify the existing mechanism, however, to the best of my knowledge, there does not exist strong evidence to support these mechanisms.

$$h_{yt} = A_t \int_0^{n_{yt}} ndF(n) \quad \text{and} \quad \ell_{yt}(a) = F(n_{yt}). \tag{C.7}$$

The formulation takes advantage of homogeneous adoption costs, which guarantee that the share of adopters is identical across available technologies.[16]

Consider the old generation next. A crucial difference is that they have already adopted technologies in the previous period for which they do not need to pay adoption costs again. Thus, old workers will only have to pay adoption costs for old technologies if they haven't learned about the technology yet, i.e. if the adoption threshold exceeds its counterpart from the previous period. For new technologies, on the other hand, old workers have to pay the full adoption costs. Again, the benefits of adopting a technology are independent of its invention date, such that the worker can simply set an adoption threshold n_{ot} with the associated costs h_{ot}:

$$h_{ot} = A_{t-1} \int_0^{n_{ot}} \mathbb{1}\{n_{yt-1} < n\} ndF(n) + a_t \int_0^{n_{ot}} ndF(n). \tag{C.8}$$

Note that the indicator guarantees that the technology has not been previously adopted by the generation. The associated labor supply then depends on the invention period as well. In particular, the adoption threshold for old technologies is the maximum of the previous period's adoption threshold and the current period's threshold. The adoption of new technologies is as in the baseline case for the young.

$$\ell_{ot}(a) = \begin{cases} F(\max\{n_{yt-1}, n_{ot}\}) & \text{if} \quad a \in A_{t-1} \\ F(n_{ot}) & \text{if} \quad a \in a_t. \end{cases} \tag{C.9}$$

Total technology adoption costs are the aggregate across generations:

$$h_t = s_y h_{yt} + (1 - s_y) h_{ot}. \tag{C.10}$$

In summary, the representative household makes technology adoption choices weighing current cost against current and future benefits, where the latter depend on wages to be earned from a particular technology. This naturally brings us to the production sector.

[16] If instead learning costs were identical across workers, optimal adoption would imply an all-or-nothing pattern for each technology without affecting the model's core predictions.

Final production. The final good y_t is produced by a representative firm using labor $\ell_t(a)$ in conjunction with equipment $k_t(a)$ for $a \in A_t$. Each technology is associated with a unique type of equipment.[17]

$$y_t = \int_{A_t} \ell_t(a)^{1-\alpha} k_t(a)^{\alpha} da. \tag{C.11}$$

The final good producer takes equipment prices $P_t(a)$ and wages $W_t(a)$ as given and solves its standard profit maximization problem:

$$\max \quad y_t - \int_{A_t} W_t(a)\ell_t(a)da - \int_{A_t} P_t(a)k_t(a)da \qquad \text{s.t.} \quad (\text{C.11}). \tag{C.12}$$

Equipment manufacturers. The blueprint for each technology is owned by an independent monopolist, who produces the associated capital good at constant marginal costs ψ in terms of the final good and sells it to the final producer at cost $P_t(a)$. To simplify the exposition I will assume that equipment fully depreciates each period. This assumption can easily be relaxed without changing any of the main results below.

Given full depreciation and market clearing, the equipment produced is the same as the equipment used and I will use the same notation. The monopolist takes into account its price effect on the demand by the final goods producer, but not the associated second-order effects on technology adoption by workers. This ensures that the analysis remains tractable. Resulting, the monopolist solves the static problem

$$\max P_t(a)k_t(a) - \psi k_t(a), \quad \text{s.t.} \quad P_t(a) = \alpha \left(\frac{\ell_t(a)}{k_t(a)}\right)^{1-\alpha}. \tag{C.13}$$

Innovation. The innovation sector is the key driver of economic growth by creating new technologies. The sector invest per capita resources x_t to generate new varieties

[17] Note that the standard expanding variety model is a special case of this production function, where all workers know about all technologies. In that case, $\ell_t(a) = 1$ and thus the production function simplifies to

$$y_t = \int_{A_t} k_t(a)^{\alpha} da.$$

a_{t+1} according to the simple linear production function:[18]

$$a_{t+1} = \varphi_0 x_t. \tag{C.14}$$

To simplify the exposition, I will directly assume that the innovation sector is governed by two equations. Firstly, equation (C.15) states the benefits of innovation per dollar invested have to be equal to the opportunity cost of investment, which is the economy's effective discount rate:[19]

$$\varphi_0 v^0_{t+1} = \left(\frac{1+r_{t+1}}{1+n}\right), \tag{C.15}$$

where v^0_{t+1} is the expected net present value of profits from a new invention and φ_0 the research productivity. Appendix C.4 shows that this can be motivated by a competitive innovation sector borrowing from the household to finance its innovation expenditures.

Secondly, the innovation sector distributes all profits to the bondholders in the economy, such that

$$r_t b_t = \int_{A_t} \pi_t(a) da. \tag{C.16}$$

The full distribution of income to bondholders can be motivated by assuming that the innovation sector does not have any equity initially and operates in perfect competition or with free entry. Due to the linear production function, this will imply zero profits and thus all income is paid to the lenders.

C.1.1 Market-clearing conditions

Finally, the economy is subject to two market-clearing conditions. Goods market-clearing requires that resources are either invested in learning, capital goods, and innovation or consumed.

[18] Formulating the production function in per capita terms neutralizes strong market size effects from population growth (see e.g. Jones (1995)). This simplifies the exposition greatly and allows me to focus on balanced growth path differences. The main results will still be in effect in a semi-endogenous growth setup, however, they will apply to the transition path of the economy instead of the balanced growth path. This is unlikely to change the short to medium term implications of the framework developed in this paper.

[19] Population growth appears in this equation as profits scale with the population size.

$$y_t = \int_{A_t} \psi k_t(a) da + h_t + x_t + c_t. \tag{C.17}$$

Secondly, market clearing in the investment sector requires that savings equal investment in innovation:

$$x_t = (1+n)b_{t+1} - b_t. \tag{C.18}$$

C.1.2 Equilibrium

I next define a competitive equilibrium in this economy and a balanced growth path equilibrium. I will focus on the latter only in my analysis below.

Definition 1. *Given $\{A_0, a_0, n_{y-1}\}$, a Competitive Equilibrium is a sequence*

$$\{y_t, h_t, x_t, c_t, A_t, a_t, n_{yt}, n_{ot}, \{k_t(a), \ell_{yt}(a), \ell_{ot}(a), \ell_t(a), P_t(a), W_t(a)\}_{a \in A_t}, r_t\}_{t=0}^{\infty}$$

such that

(a) the representative household, the final good producer, and the producers of intermediate goods solve their maximization problems,

(b) the no-arbitrage condition in the investment sector holds,

(c) markets clear.

Definition 2. *A Balanced Growth Path is a competitive equilibrium such that consumption grows at constant rate g.*

C.2 Equilibrium Characterization

I will limit the equilibrium characterization to the core results that are necessary to understand the intuition of the model. Detailed derivations and proofs are provided in Appendix Section C.5.

Lemma C.1. *On any BGP, the interest rate satisfies $1 + r = \frac{1+g}{\beta}$. Furthermore, as long as $g \geq 0$, the effective discount rate of the economy satisfies $\frac{1+r}{1+n} > 1$.*

Technology adoption and wages. To simplify the analysis and abstract from corner solutions, I will assume that adoption cost follow a continuous distribution with unbounded support from above.

Assumption 1. *The cost distribution function satisfies $f(n) > 0$ for $n \in (0, \infty)$, where $f(n)$ is the pdf of $F(n)$.*

Lemma C.2. *On any BGP, tasks wages $\mathcal{W}$ are constant and identical across tasks. Furthermore, the adoption thresholds for young and old workers are constant over time and given by*

$$n_y = \mathcal{W}\left(1 + \frac{1-p}{1+r}\right) \quad \textit{and} \quad n_o = \mathcal{W}. \tag{C.19}$$

Firstly, note that constant wages per variety are a standard result in expanding variety models with constant marginal costs of production in the intermediary sector. In particular, the capital-labor ratios in the model, which determine the wages, are directly linked to the equilibrium price of the intermediary good, which in turn is supplied at a constant markup over marginal costs. Since the latter is constant and identical across equipment varieties, wages are as well.

The second part of the Lemma is a direct result of the first. As all technologies yield the same benefits, workers only differentiate between them according to their adoption costs. The benefits of adoption are then the expected, discounted wages earnings. The marginal adopted technology type equalizes cost and benefits. For the old generation, this implies that all technologies yielding weakly positive net income are adopted, while the young generation adopts technologies whose current and future expected, discounted benefits exceed current adoption costs.

Corollary C.1.

(a) Workers adopt technologies as early as possible or never.

(b) Old workers have lower technology adoption rates driven by threshold differences for new technologies.

(c) Take-home income is increasing in age over the life cycle and in the cross-section.

(d) *Old technologies have higher aggregate technology adoption rates than young technologies.*

Consider (a) first. The payoff from learning about a technology is strictly increasing in the number of periods that a given generation can use it in the labor market, while the adoption costs stay constant. Thus, it is always preferable to adopt a technology early if ever.

Part (b) links the insight of early adoption to differences in the availability of technologies over time. In particular, old workers adopted old technologies when they were young and, thus, due to the constant adoption threshold for each age group, young and old workers adopt the same share of old technologies. In contrast, old workers apply their current, lower adoption threshold to new technologies as they did not have the opportunity to learn about them previously. Via a simple composition effect across old and new technologies, this implies that old workers have lower aggregate technology adoption rates compared to young workers, who apply the same, high technology adoption threshold to all currently available technologies.

Note that higher aggregate technology adoption rates also imply larger skill sets for young workers. The latter might be perceived as a bug rather than a feature given the extensive evidence for increasing compensation over the life-cycle (See e.g. Lagakos et al. (2018)). While the model does not possess features that are likely important for life-cycle wage dynamics such as job-ladders or learning-by-doing, it still features an upwards sloping take-home income, which I define as gross income minus adoption costs, in cross-section and across the life-cycle as pointed out in part (c).

Two insights are driving this result. Firstly, old workers gain more from old technologies as they do not have to pay their adoption costs again. Secondly, old workers also gain more from new technologies as they adopt all new technologies that generate positive net cash flow in this period. On the other hand, young workers adopt some technologies with negative cash flow in the current period due to the benefits in the next period. As a result, old workers receive larger take-home income from the labor market.

Finally, and as pointed out in (d), technologies themselves are subject to a life-cycle pattern, which arise due to composition effects. Over time, low adoption generations, i.e. the initially old, are replaced by high adoption generations. Eventually, all active generations entered the economy when the technology was available and, thus, had the chance to adopt it when young. Therefore, for a given technology, the aggregate adoption rate has an upwards trajectory converging towards its long-run value, the adoption rate of young workers.

Firm Profits and the Value of Innovation. Having solved the worker problem, we can next turn our attention to the intermediary problem.

Lemma C.3. *Per capita profits for a variety are proportional to its adoption rate:*

$$\pi_t(a) = \tilde{\pi}\ell_t(a). \tag{C.20}$$

Similarly, the per capita value of a new variety is proportional to its discounted market size:

$$v^0 = \tilde{\pi}\left(\ell^N + \left(\frac{1+n}{r-n}\right)\ell^E\right), \tag{C.21}$$

where $\ell^N = s_y F(n_y) + (1-s_y)F(n_o)$ and $\ell^E = F(n_y)$ are the aggregate technology adoption rates for new and old technologies respectively.

Firstly, note that the formulation for profits is standard in the endogenous growth literature apart from the explicit acknowledgment of adoption rates as a driver of market size. The latter matter for per capita profits as the monopolist earns constant profits per adopter.

Market size effects for profits directly bleed into the value of a new innovation. The key insight from is formulation is that the adoption rate for new technologies only matters in the first active period as the technology becomes an old technology afterward. Note that the expansion of market size for old technologies is directly linked to the fact that they are adopted by young workers only. As a result, the workforce age composition matters for short-run profits, but not in the long run.

How does aging impact the model economy? Before understanding the effects of aging in the model, I quickly note that the BGP exists and is unique.

Proposition C.1. *There exists a unique balanced growth path equilibrium.*

To gain some insight into the model dynamics I will discuss a set of comparative statics exercises. I start by taking the WYS s_y as exogenous in partial equilibrium and then discuss how the intuitions developed for this simple scenario translate to general equilibrium.

Proposition C.2. *Holding the constant the interest rate and population growth rate, an exogenous decline in the WYS decreases the average adoption rate for new and overall technologies, (gross) output, and the value of new inventions.*

The important insight is that there are pure composition effects from the WYS pushing down technology adoption, output, and the value of new innovations. The next proposition highlights how these feed into general equilibrium.

Proposition C.3. *Holding constant the population growth rate, an exogenous decline in the WYS decreases the aggregate adoption rate for new and overall technologies, investment into new technologies relative to old technologies, the value of new inventions, the interest rate, and the economy's productivity growth rate.*

The key insight from the proposition is that the partial equilibrium results based on Proposition C.2 carry over into general equilibrium. In response to declining firm values, interest rates have to decline as well to satisfy the research arbitrage equation. Lower interest rates translate to lower productivity growth rates via the Euler equation. The overall mechanism is clear: Population aging reduces the technology adoption rate for new innovations via a simple composition effect. Declining adoption rates decrease the value of innovation and, thus, lead to a reduction in R&D investment. The resulting decline in innovation directly implies lower productivity growth rates.

Finally, the next proposition confirms that these predictions carry over to a decline in the working young share driven by declining population growth rates, which is the empirically relevant case for the US. The decline in fertility itself has first-order consequences via market size effects, which turn out to point in the same direction as the composition effects.

Proposition C.4. *A decrease in the population growth rate, which mechanically leads to a decrease in the WYS, decreases the aggregate adoption rate for new and overall* *technologies, investment into new technologies relative to old technologies, the value of new inventions, the interest rate, and the economy's productivity growth rate.*

What are the policy implications of an aging economy? Given the results above, the question arises of whether there is room for policy in this framework. To study this question, I introduce the social planner problem in Appendix C.3 and focus on its implications here:

Proposition C.5. *The social planner solution features higher technology adoption rates for older workers, a flatter life-cycle profile of adoption thresholds, and a higher productivity growth rate.*

Inefficiently low productivity growth rates are a ubiquitous feature of the endogenous growth literature as firms are unable to capture the full value of their innovation, e.g. because part of it is paid to workers in wages. Similarly, monopoly distortions feed into inefficiently low wages, which, in this framework, translate into inefficiently low adoption rates. Setting optimal capital-labor ratios immediately yields higher adoption rates. The adoption profile flattens as future resources generated by young workers are discounted at a higher rate due to faster economic growth, providing a countervailing force for young workers to the overall larger marginal product of technology adoption. Since old workers do not have future income, they are only subject to the pure increase in marginal product effect.

Proposition C.6. *In the Social Planner Equilibrium, a decrease in the population growth rate, which mechanically leads to a decrease in the WYS, decreases the aggregate technology adoption rate as well as the economy's productivity growth rate.*

Proposition C.6 is the social planner equivalent to Proposition C.4 and highlights that the direction of the response to an aging population is the same across solution concepts. Thus, while adoption levels and innovation activity are sub-optimally low in the competitive equilibrium, its response to an aging population is not necessarily sub-optimal. The intuition for this result is that the forces leading to a declining productivity growth rate in the competitive equilibrium are still active in the social planner solution. Lower population growth rates lower the value of resources in the future. Furthermore, adoption rates decline as well due to changes in the relative weight of resources across periods, leading to a declining social value of innovation as well. Thus, while adoption levels and innovation activity are sub-optimally low in the competitive equilibrium, their responses to an aging population are not necessarily sub-optimal.

C.3 Social Planner Solution

Decision problem. The equations for the planner setup are provided below. I forgo proving that $n_y > n_o$ in equilibrium and directly impose it here. This is without loss of generality as there are no inefficiencies in the adoption conditional on factor rewards.

$$\begin{aligned}
&\max \sum_{s=0}^{\infty} \beta^s (1+n)^s \ln(c_{t+s}), \\
\text{s.t.} \quad &\int_{A_t} \ell_t(a)^{1-\alpha} k_t(a)^{\alpha} da = \int_{A_t} \psi k_t(a) da + h_t + x_t + c_t \\
&\ell_t(a) = \begin{cases} s_y F(n_{yt}) + (1-s_{yt}) F(n_{yt-1}) & \text{if} \quad a \in A_{t-1} \\ s_y F(n_{yt}) + (1-s_{yt}) F(n_{ot}) & \text{if} \quad a \in a_{t-1}. \end{cases} \\
&h_t = s_y A_t \int_0^{n_{yt}} n dF(n) + (1-s_y) a_t \int_0^{n_{ot}} n dF(n) \\
&A_{t+1} = A_t + \varphi_0 x_t
\end{aligned}$$

Naturally, we have to add the appropriate initial conditions on technology and previous adoption.

Definition 3. *A social planner equilibrium is a set of sequences*

$$\{y_t, h_t, x_t, c_t, A_t, a_t, n_{yt}, n_{ot}, \{k_t(a), \ell_{yt}(a), \ell_{ot}(a), \ell_t(a)\}_{a \in A_t}\}_{t=0}^{\infty}$$

such that the social planner maximizes its objective functions subject to its constraints and markets clear.

Definition 4. *A Balanced Growth Path for the social planner problem is a social planner equilibrium such that consumption grows at constant rate g.*

C.4 Motivating the Investment Sector

I briefly outline a investment sector problem that gives rise to the equations presented in the text. There is a representative investment firm producing new innovations with production function

$$a_{t+1} = \varphi_0 x_t.$$

To finance innovation, the firm borrows from the households at rate r_t such that its (discounted) profits from new investments are given by

$$\left(\frac{1+n}{1+r_t}\right) \mathbb{E}_t \left[\int_0^{a_{t+1}} v_{t+1}^0(a) da\right] - x_t,$$

where $v_{t+1}^0(a)$ is the value of new innovation a at time $t+1$, which equals the present discounted value of profits. Due to the linearity of the investment function and homogeneous adoption costs, it follows immediately that any interior solution needs to satisfy

$$\varphi_0 \left(\frac{1+n}{1+r_t}\right) v_{t+1}^0 = 1,$$

The second equation can be motivated by assuming that the sector is fully leveraged at $t = 0$. From the equation it follows immediately that the sector never builds equity such that $r_t b_t$ has to equal all the profits earned by the sector.

C.5 Derivations and Proofs

This section provides derivations and proofs omitted from the main text.

Production and Prices. First order condition for the final producer's problem yield the standard factor demands:

$$P_t(a) = \alpha \left(\frac{k_t(a)}{\ell_t(a)}\right)^{\alpha-1} \qquad \text{and} \qquad W_t(a) = (1-\alpha)\left(\frac{k_t(a)}{\ell_t(a)}\right)^{\alpha}.$$

Monopolist solves the profit maximization problem taking into account the equipment demand for monopolist price $P_t(a)$, which in turn pins down the equilibrium capital-labor ratio $\mathcal{K}$ and equilibrium task wage $\mathcal{W}$ via the first order conditions of the final good producer:

$$P_t(a) = \mathcal{P} = \frac{\psi}{\alpha}, \quad \frac{k_t(a)}{\ell_t(a)} = \mathcal{K} \equiv \left(\frac{\mathcal{P}}{\alpha}\right)^{\frac{1}{\alpha-1}}, \quad \text{and} \quad W_t(a) = \mathcal{W} \equiv (1-\alpha)\mathcal{K}^{\alpha}.$$

Plugging in the definition of $\mathcal{K}$ and $\mathcal{P}$ yields the expression in Lemma 1 for the task wage.[20] Note that we can already solve for firm profits nd the value of a new invention conditional on household adoption:

$$\pi_t(a) = (P_t(a) - \psi)k_t(a) = (1-\alpha)\alpha^{\frac{1}{1-\alpha}}\mathcal{P}^{-\frac{\alpha}{1-\alpha}}\ell_t(a) = \alpha\mathcal{W}\ell_t(a).$$

The value of a new invention is then just the expected, discounted value of profits.

Household Decisions. With the skill wages in hand we can turn our attention to the household problem.

Proof of Lemma C.1. Note that this is the standard Euler equation result. In particular, the first order conditions of the household for b_{t+1} and c_t require

$$1 = \beta(1+r_t)\frac{c_t}{c_{t+1}}.$$

By definition of a BGP $c_t/c_{t+1} = 1/(1+g)$ and the first result follows. The second part follows by rearranging the Euler equations and noting that $\beta(1+n) < 1$. by assumption. Thus, as long as $g \geq 0$, we have effective discounting. $\square$

Consider next the first order conditions for the adoption threshold of old workers. This does not have any inter-temporal implications and thus simply involves maximizing the net-resources for the household:

[20] Note that $\partial\mathcal{W}/\partial\mathcal{P} < 0$, i.e. the equipment price set by the intermediary producer reduces the task wages via its impact on the capital-labor ratio. This will become important once we consider adoption rates by households. In particular, it will be the case that adoption is increasing in the tasks wage. As a result, the intermediary producer has an incentive to decrease prices as to increase the market size. I will abstract from this consideration, but note that this will naturally lead to a lower markup compared to the case considered here, but higher profits. Allowing the intermediary producers to take into account this impact makes the problem intractable.

$$(1 - s_y) f(n_{ot}) \mathcal{W} = (1 - s_y) f(n_{ot}) n_{ot}.$$

The left hand side states the gross resources generated at the margin, which is the mass of workers times the mass of technologies at the threshold times the (constant) task wage. This has to be equal the cost at the margin, which are the mass of workers to which the threshold applies time the mass of technologies at the threshold (since the household has to pay for all of them) time the cost per technology at the threshold, which is the threshold itself. Following the assumption that $f(n_{ot}) > 0$, the condition simplifies to the constant adoption threshold in the text. Positive support ensures the the threshold is clearly defined and unique. Having $f(n) = 0$ for some n potentially gives rise to saddle points or sets of optimal thresholds.

Note that I've implicitly assumed that the marginal value of resources is positive and have already normalized by the mass of technologies around the threshold, which could be a_t or A_{t-1} given the threshold. Both terms will show up on both sides and thus do not influence the adoption threshold.

Next, consider the problem for choosing the adoption threshold for the young household. I will first take the derivative assuming that $n_{yt} > n_{ot}$ and then confirm this conjecture. Furthermore, I will highlight that assuming the opposite does not yield a solution in line with the conjecture.

The first order condition for s_{yt} can be derived as

$$s_y f(n_{yt}) f(n_{yt}) \mathcal{W} + \frac{\lambda_{t+1}}{\lambda_t} (1 - s_y) f(n_{yt}) \mathcal{W} = s_y f(n_{yt}) n_{yt}.$$

Firstly, note that the right hand side is the same as before. Secondly, consider the LHS. The first term is as for the old generation and represents current gains. The second term represents future gains from current adoption, appropriately discounted by the relative value of resources λ_{t+1}/λ_t, where λ_t is the Lagrange multiplier on the resource constraint. Furthermore, note that mortality risk is taken into account as the benefits only apply to a mass $(1 - s_y)$ of workers.

Plugging the Euler condition for the relative value of resources across periods and normalizing by s_y yields the expression for n_y in the text. Note that the expression satisfies $n_y > n_o$ as per our conjecture.

Now instead suppose $n_{yt} < n_{ot}$. Then the resulting first order derivative can be expressed as

$$s_y f(n_{yt}) f(n_{yt}) \mathcal{W} = s_y f(n_{yt}) n_{yt} - \frac{\lambda_{t+1}}{\lambda_t}(1 - s_y) f(n_{yt}) n_{yt}.$$

Firstly, note that the benefit are only current period, as the future adoption threshold being larger than the current one implies that the technology will be adopted tomorrow anyways and thus tomorrows benefits do not depend on today's action. On the other hand, the cost of adoption reflect both current period adoption costs as well as the savings made next period. In particular, adopting the technology today implies that the household does not have to pay for the adoption tomorrow. It is straight-forward to show that the associated adoption threshold with this first order condition violates $n_o > n_y$ and thus this can never be an equilibrium.

Proof of Lemma C.2. See derivations above. □

Proof of Corollary C.1. See derivations above for part (a).

For part (b) note that it follows immediately from (C.19) that young workers adopt new technologies at a higher rate. In particular, the adoption rate for new technologies for either generation is $F(n_y)$ and $F(n_o)$ respectively. Given that $n_y > n_o$ and $F(\cdot)$ is a strictly increasing function, the latter will always be larger. This carries over to the overall adoption rate via a simple composition effect. The share of adopted technologies among A_t for each age group, denoted by $\mathcal{A}^y$ and $\mathcal{A}^o$ respectively, is given by:

$$\mathcal{A}_y = \frac{A_t F(n_y)}{A_t} = F(n_y) \quad \text{and} \quad \mathcal{A}_o = \frac{A_{t-1} F(n_y) + a_t F(n_o)}{A_t} = \frac{1}{1+g} F(n_y) + \frac{g}{1+g} F(n_o).$$

Given that $n_o < n_y$, it follows immediately that $\mathcal{A}_y > \mathcal{A}_o$ for $g > 0$.

Next, consider part (c). The proof for the first part of this is straight-forward when considering the net income earned by a young worker. In particular, let w_{yt} the gross income of the young generation, then we can decompose the overall net income as

$$\begin{aligned}
w_{yt} - h_{yt} =& A_t \int_0^{n_y} (\mathcal{W} - n) dF(n) \\
=& A_{t-1} \int_0^{n_y} (\mathcal{W} - n)\, dF(n) + a_t \int_0^{n_o} (\mathcal{W} - n)\, dF(n) + a_t \int_{n_o}^{n_y} (\mathcal{W} - n)\, dF(n)
\end{aligned}$$

The first line states that the net income for young workers is the mass of available technologies times the integral over the net benefits from each adopted technology type. The second line splits this into the net benefits for technologies that the old generation adopted when young plus the net benefits of the new technologies adopted by the old in the current period plus the net benefits from new technologies adopted by the young, but not by the old. We can compare this to the same calculation for old workers:

$$w_{ot} - h_{ot} = A_{t-1} \int_0^{n_y} \mathcal{W} dF(n) + a_t \int_0^{n_o} (\mathcal{W} - n)\, dF(n).$$

Note that old workers do not have to pay the adoption cost for technologies adopted when they were young. The comparison across terms is quite straight-forward then. Old workers have a clear advantage in the first terms. The second term is the same for both and, finally, the third term for young workers is always negative. One can show this immediately by noting that $\mathcal{W} - n_o = 0$ by definition of the adoption threshold. Thus $\mathcal{W} - n$ is going to be negative for all $n > n_o$. The intuition is straight-forward. Old workers adopt all technologies that help them in the present. Thus, if there is a technology that young adopt, but old do not, then this technology cannot yield positive returns in the present. Note that the present discounted value is still going to be positive from the future income flow.

For the second part, note that the we can express the income of an old generation tomorrow as

$$w_{ot+1} - h_{ot+1} = A_t \int_0^{n_y} \mathcal{W} dF(n) + a_{t+1} \int_0^{n_o} (\mathcal{W} - n) dF(n).$$

It is trivial to show that this exceeds $w_{yt} - h_{yt}$.

Finally, for part (d) note that old technologies, i.e. technologies invented in the previous period, were adopted by the current old generation when they were young.

Furthermore, the current adopters are the young generation as well. This yields an economy with adoption rate of $F(n_y)$. In contrast, new inventions are first adopted by the current new and old generations. As a result, their adoption rate is simply $s_y F(n_y) + (1 - s_y)F(n_o)$. Given that $n_y > n_o$, this is smaller than $F(n_y)$. $\square$

The Value of New Innovations. Having determined technology adoption rates, we can turn our attention back to the value of innovation. Note that an invention is a new technology in its first period and an old afterwards. Thus, $\ell_t(a) = s_y F(n_y) + (1 - s_y)F(n_o)$ in its first period and $F(n_y)$ in all following periods. Thus, the (per capita) value of a new invention is given by

$$\begin{aligned} v^0 &= \sum_{s=0}^{\infty} \left(\frac{1+n}{1+r}\right)^s \mathbb{E}[\pi_{t+s}(a) | a \in a_t] \\ &= \alpha \mathcal{W} \left(s_y F(n_y) + (1 - s_y)F(n_o) + \sum_{s=1}^{\infty} \left(\frac{1+n}{1+r}\right)^s F(n_y) \right). \end{aligned}$$

Note that $(1+n)^s$ corrects for population growth. The formula in the text simply solves the infinite sum and rearranges terms.

Furthermore, note that by a similar calculation, we can determine the value of old technologies as

$$v^E = \alpha \mathcal{W} \left(\frac{1+r}{r-n}\right) F(n_y).$$

The only difference being that the adoption rate is constant for all periods.

Lemma C.4. *There exists a unique interest rate r that satisfies the research arbitrage equation. Furthermore, there exist $\underline{\varphi}_0$ such that $\forall \varphi_0 \geq \underline{\varphi}_0$, the equilibrium growth rate satisfies $g \geq 0$.*

Proof of Lemma C.4. Firstly, we can use our results in the previous lemmas to rearrange the research arbitrage equation to

$$\frac{1+n}{1+r} v_0 = \frac{1}{\varphi_0}.$$

Note that the RHS is constant in r. The LHS, in contrast, is strictly decreasing in r for two reasons. Firstly, and increase in r increases the discount rate, which lower the value of future profits. Since all terms are discounted, this has a strictly negative effect. Secondly, an increase in r also pushes down n_y, which further decreases the value of innovation. Given that all these effects are strict and point in the same direction, we have a strictly decreasing function in r on the LHS. In other words, if there exists an interest rate satisfying this condition, then it is unique.

To show existence, note that $\lim_{r\to n}\left(\frac{1+n}{1+r}v_0\right) \to \infty$ and $\lim_{r\to\infty}\left(\frac{1+n}{1+r}v_0\right) \to 0$. Thus, as long as $\varphi_0 \in (0,\infty)$, there exists an $r > n$ to satisfy this equation.

For the second part, note that since the LHS is decreasing in r and the RHS is decreasing in φ_0, there exist and implicit function $r(\varphi_0)$ that is strictly increasing in φ_0. We can then take advantage of Lemma 1 stating that

$$1 + g = \beta(1 + r(\varphi_0)),$$

to note that $\exists \underline{\varphi}_0$ such that $\beta(1 + r(\varphi_0)) > 1\ \forall \varphi_0 > \underline{\varphi}_0$. □

Aggregates and Market Clearing. The no profit condition in the innovation sector as well as market clearing for savings imply a simplified budget constraint for households:

$$w_t + \pi_t = c_t + h_t + x_t,$$

where π_t denoted the aggregate profits. Note that $w_t + \pi_t = y_t - i_t$. Furthermore, by the research production function, we have $x_t = a_{t+1}/\varphi_0$. Denote by $\tilde{y} = y_t/A_t$ with similar definitions for other variables, then we can rearrange the resource constraint to

$$\tilde{y} = \tilde{c} + \tilde{i} + \tilde{h} + \frac{g}{\varphi_0}.$$

It is straight-forward to be shown that $\tilde{c} > 0$ on the balanced growth path. Furthermore, one can show that $\lim_{s\to\infty} \lambda_{t+s} = 0$ as $\lambda_{t+s} = \left(\frac{1+n}{1+r}\right)^s \lambda_t$, $\lambda_t > 0$ and $r > n$. Thus, the problem is well defined.

Finally, note that for any other balanced growth path equilibrium we have

$$\tilde{\lambda}_{t+s} = \tilde{\lambda}_t \left(\frac{1+n}{1+r}\right)^s = \tilde{\lambda}_t \left(\frac{\beta(1+n)}{1+g}\right)^s \tag{C.22}$$

By assumption (via $\varphi_0 \geq \underline{\varphi}_0$ and $x_t \geq 0$), we have $g \geq 0$. Since $\beta(1+n) < 0$ and $\tilde{\lambda}_t \geq 0$ (from $c_t \geq 0$), we have $\lim_{s\to\infty} \tilde{\lambda}_t \left(\frac{\beta(1+n)}{1+g}\right)^s \in (0, \infty)$. Thus, all other balanced growth path solutions are also well defined.

Main Results.

Proof of Lemma C.1. Firstly, note that Lemma C.4 shoes that there always exists and interest rate and thus a growth rate to satisfy the research arbitrage equation. I will focus on the case with a interest rate implying a positive growth rate here.

The derivations above further show that the balanced growth path constructed so far features positive consumption and thus is optimal among balanced growth paths with bounded utility.

What remains to be shown then is that the objective function is well defined on any balanced growth path. This is straight-forward. On a BGP we have $c_{t+s} = c_t(1+g)^s$, and thus

$$\sum_{s=0}^{\infty}((1+n)\beta)^s \ln(c_{t+s}) = \ln(c_t)\sum_{s=0}^{\infty}((1+n)\beta)^s + \ln(1+g)\sum_{s=0}^{\infty}((1+n)\beta)^s s.$$

It is straight-forward to show that both terms are well defined and bounded for any $g \geq 0$. Thus, the objective function is well defined for any BGP equilibrium. This in turn implies that the equilibrium defined in the derivations above is as a matter of fact unique. Note that uniqueness follows from a unique r and thus g satisfying the research arbitrage equation. □

Proof of Proposition C.2. The proposition highlights the pure composition effects from an increase in the young share. The proof simply relies on $n_y > n_o$ and is omitted for brevity. Note that the output result follows from the fact that output is proportional to the average technology adoption rate. □

Proof of Proposition C.3. I will start the proof from the last point. Consider the research arbitrage equation:

$$\frac{1+n}{1+r}\alpha\mathcal{W}\left[\left(\frac{1+r}{r-n}\right)F(n_y)+(s_y-1)\left(F(n_y)-F(n_o)\right)\right]=\frac{1}{\varphi_0}.$$

It is straight-forward to show that an increase in s_y increases the LHS holding everything else equal, while leaving the RHS untouched. The only variable on the LHS that can respond to keep the equality is r. As per our earlier discussion, the LHS is strictly decreasing in r, thus we have that an increase in s_y needs to be offset by an increase in r. Furthermore, from the Euler equation, we know that an increase in r requires an increase in g, which completes the proof for the last bullet point.

For the third bullet point, note that since $\frac{1+r}{1+n}v_0$ is constant, but r is increasing, we need to have v_0 increasing in s_y.

The first and second bullet point are tightly linked. Let $\ell^N = s_y F(n_y) + (1-s_y)F(n_o)$ and $\ell^E = F(n_y)$ be the economy wide adoption rates of new and old technologies respectively. We can express the value of a new innovation as

$$v^0 = \alpha\mathcal{W}\left(\ell^N + \sum_{s=1}^{\infty}\left(\frac{1+n}{1+r}\right)^s \ell^E\right)$$

From before, we know that $\partial v^0/\partial s^y > 0$. Furthermore, we know that $\partial r/\partial s_y > 0$ and thus $\partial \ell^E/\partial s_y < 0$. Thus, the only way to have $\partial v_0/\partial s_y > 0$ is $\partial \ell^N/\partial s_y > 0$. In other words, the direct effect has to be stronger than the general equilibrium force pushing against it. This proves the first bullet point.

Finally, the ration of investment in new technologies to investment in old technologies can be expressed as

$$\frac{\int_{a_t}\psi k_t(a)da}{\int_{A_{t-1}}\psi k_t(a)da} = \frac{g}{1+g}\frac{\ell^N}{\ell^E}$$

Since both factors are increasing in s_y, the overall term is as well. Note that total investment in new technologies, $a_t\ell^N\mathcal{K}$, is increasing in s_y as well. □

Proof of Proposition C.4. The proof for this follows the same steps as above and is omitted for brevity. Note, however, that the induced increase in r is larger as there are two channels at play in the innovation sector: Pure market size via population growth and composition changes via s_y. □

Social planner results. Throughout this section I will omit most of the algebraic intermediate steps for brevity. Detailed derivations are available upon request.

Firstly, note that the social planner will set a higher capital-labor ratio compared to the competitive solution due to the lack of monopoly pricing.

Lemma C.5. *On a social planner BGP, the social planner chooses a higher capital-labor ratio $\mathcal{K}^{SP}$ compared to the competitive equilibrium, which implies a higher implicit wage $\mathcal{W}^{SP}$. Furthermore, the planner chooses larger technology adoption threshold n_y^{SP} and n_o^{SP} compared to the competitive equilibrium due to larger implicit wage/ the larger marginal product of labor.*

Proof. Firstly, note that the standard first order conditions for capital imply

$$\frac{k_t(a)}{\ell_t(a)} = \mathcal{K}^{SP} \equiv \left(\frac{\psi}{\alpha}\right)^{-\frac{1}{1-\alpha}}.$$

Since $\alpha < 1$, we have $\mathcal{K}^{SP} > \mathcal{K}$. This is a direct implication of the monopoly friction. The monopolist reduces supply to maximize profits, while the planner chooses the social optimum. As a direct implication of lower capital-labor ratios, we have that the implicit wage or marginal product of labor is larger in the social planner solution

$$\frac{\partial y_t}{\partial \ell_t(a)} = \mathcal{W}^{SP} \equiv (1-\alpha)\left(\frac{\psi}{\alpha}\right)^{-\frac{\alpha}{1-\alpha}}$$

Again, it is straight-forward to see that since $\alpha < 1$, $\mathcal{W}^{SP} > \mathcal{W}$. This is important since it directly impacts optimal technology adoption. In particular, we have

$$n_y^{SP} = \mathcal{W}^{SP}\left(1 + \frac{\beta(1-p)}{1+g}\right) \quad \text{and} \quad n_o^{SP} = \mathcal{W}^{SP}$$

Note that $n_o^{SP} > n_o$ in general, while $n_y^{SP} > n_y$ conditional on g. It remains to be shown whether this will be the case once we endogenize g. Furthermore, note that we can make this comparison by plugging in the Euler equation for the competitive equilibrium in n_y. □

Lemma C.6. *The social planner chooses a higher equilibrium growth rate g^{SP} compared to the competitive solution.*

Proof. It is useful to make a couple of definitions first. Denote by λ_t^{SP} the Lagrange multiplier on the resource constraint. Furthermore, denote by ℓ_N and h_N the adoption rate and associated learning costs for a new variaty and by ℓ_E and h_E the associated values for existing varieties. One can then show that the first order conditions for x_t boil downs to

$$\frac{1}{\varphi_0} = \frac{\lambda_{t+1}}{\lambda_t}\left(\mathcal{W}^{SP}\ell_N - h_N\right) + \sum_{s=2}^{\infty}\frac{\lambda_{t+s}}{\lambda_t}\left(\mathcal{W}^{SP}\ell_E - h_E\right)$$

Note that the LHS denotes the unit costs of innovation, while the RHS denotes the benefits discounted to current marginal utility. These benefits are the net-gains from a new technology tomorrow plus the net-gains of an old technology starting in two periods. Note that investment costs are already taken into account in this formulation.

Plugging in the evolution of marginal products along the BGP, we have

$$\frac{1}{\varphi_0} = \frac{(1+n)\beta}{1+g}\left(\left(\mathcal{W}^{SP}\ell_N - h_N\right) + \sum_{s=1}^{\infty}\left(\frac{(1+n)\beta}{1+g}\right)^s\left(\mathcal{W}^{SP}\ell_E - h_E\right)\right)$$

Define the implicit value of innovations as

$$v_{SP}^0 = \left(\left(\mathcal{W}^{SP}\ell_N - h_N\right) + \sum_{s=1}^{\infty}\left(\frac{(1+n)\beta}{1+g}\right)^s\left(\mathcal{W}^{SP}\ell_E - h_E\right)\right).$$

Note that to show that $g^{SP} > g$, we need to show that $v_{SP}^0 > v^0$. To see why this is true, note that in the competitive market equilibrium, total generated resources from innovation are v^0 plus the net-present values of wages minus adoption costs. Note that the latter are strictly positive by the first order conditions of workers. Denote by v_P^0 the sum of both and by $v_{SP}^0(g)$ the social planner value associated with a growth rate as in the competitive equilibrium. It follows that $v^0 < v_P^0 \leq v_{SP}^0(g)$. The first inequality follows from positive net-income of workers and the second from the fact that (conditional on g), the social planner can always enact the competitive equilibrium solution. However, this implies

$$\frac{1}{\varphi_0} = \frac{(1+n)\beta}{1+g} v^0 < \frac{(1+n)\beta}{1+g} v^0_{SP}(g)$$

Note that the I've used the Euler equation for the expression for the competitive solution. Finally, since $\frac{(1+n)\beta}{1+g} v^0_{SP}(g)$ is strictly decreasing in g, the equilibrium with $\frac{1}{\varphi_0} = \frac{(1+n)\beta}{1+g} v^0_{SP}(g^{SP})$ needs to satisfy $g^{SP} > g$.

□

Proof of Proposition C.5. The proposition follows from the results above. □

Proof of Proposition C.6. To proof this result, it is convenient to rewrite the "research arbitrage equation" in terms of the resources generated for each generation:

$$\frac{1}{\varphi_0} = \left(\frac{(1+n)\beta}{1+g}\right) \left((1-s_y) \left(F(n_o) \mathcal{W}^{SP} - h_o \right) + s_y \sum_{s=0}^{\infty} \left(\frac{(1+n)\beta}{1+g}\right)^s \left(\left(1 + \frac{(1-p)\beta}{1+g}\right) \mathcal{W}^{SP} F(n_y) - h_y \right) \right)$$

From the optimal technology adoption choice if follows that

$$F(n_o)\mathcal{W}^{SP} - h_o = \int_0^{n_o} (n_o - n) dF(n) < \int_0^{n_y} (n_y - n) dF(n) = \left(1 + \frac{(1-p)\beta}{1+g}\right) \mathcal{W}^{SP} F(n_y) - h_y.$$

Thus, a decrease in s_y pushes down the right hand side and, thus, needs to be offset by a correspondingly lower growth rate. A decrease in n has the same effect and thus both forces push in the same direction.

The decline in the average technology adoption rate is due to the simple composition effect that is only partly offset by the decline in g. The proof for this is similar to the one for the competitive equilibrium and omitted here for brevity. □

D Evidence on the Age-Technology Adoption Nexus

The computer has arguably been the most important "new" production technology introduced in the 1990s and early 2000s. Earlier studies document its wide ranging impact on firm productivity and demand for skills across industries and occupations (Autor et al., 1998, 2003; Brynjolfsson et al., 2002; Bresnahan et al., 2002). Nonetheless, computer adoption was not uniform across workers and, as documented below, older workers' adoption rates significantly lagged their younger counterparts.

In this section, I carefully document that older cohorts had lower adoption rates of the computer at the workplace in the 1990s and early 2000s. The analysis expands on Friedberg (2003) by using a longer time frame, extended set of outcome variables, and a non-parametric regression approach controlling for a wider set of confounding factors such as occupation and industry choice. This evidence motivates the model developed in the subsequent section.[21]

D.1 Data

I investigate computer adoption at the workplace using the five CPS Computer and Internet (CIU) Supplement waves between 1989 and 2003. (Flood et al., 2020) I limit my analysis to responses linked to use at work to capture differences in the adoption of productive technologies. I restrict the sample to full-time employees between the age of 25 and 64 with at least a high school degree. This is intended to ensure that the computer was a relevant technology for the worker and that differences in effective labor supply are not driving my results.

I construct two measures of computer adoption by workers. Firstly, I consider a simple indicator measure of computer use at work, which I will refer to as computer adoption, which is based on the response to the question of whether the respondent uses a computer at work. Secondly, I construct a proficiency index by counting the number of tasks a worker performs with a computer at work conditional on working with it at all. The task index ranges from 1 to 6 and is only available for workers reporting computer use at work. The list of tasks performed with the computer that are consistently available throughout the survey years include calendar/scheduling, databases or spreadsheets, desktop publishing or word processing, electronic mail and programming.[22] I do not consider tasks that were not consistently asked throughout the survey waves to ensure that the estimation is not capturing changes in the survey

[21]See also Weinberg (2004); Aubert et al. (2006); Meyer (2007), and Schleife (2008) for related evidence on technology adoption across the lifecycle.

[22]"databases or spreadsheet" and "desktop publishing or word processing" are split into the individual items during the first three survey waves, but combined during the latter two. I aggregate both to have a consistent measure throughout.

structure. I will refer to this variable as the proficiency index.

Besides the CIU specific variables, I use the age and gender of the respondent, state of residency, educational attainment, occupation, and industry. I use occupational codes that are standardized using the 1990 definitions as provided by IPUMS. For industry classifications, I use the code provided on David Dorn's data page.[23]. Throughout I use 5-year year-of-birth cohorts starting from 1924-28 and report the results by transforming the cohort measure into age groups in 1989 to aid interpretation. Table D.1 reports summary statistics.

Table D.1: Summary Statistics for CPS Sample

Variable	**Obs.**	**Mean**	**Std. Dev.**
PC Adoption	207,998	0.581	0.493
PC Proficiency	109,280	2.784	1.703
Age	207,998	41.013	9.934
Female	207,998	0.423	0.494
College Degree	207,998	0.341	0.474
Graduate Degree	207,998	0.125	0.330
White	207,998	0.846	0.361
Black	207,998	0.106	0.308
Asian	207,998	0.038	0.191

Note: This tables reports summary statistics for the CPS CIU sample. Observations are weighted by CPS CIU supplement weights.

D.2 Empirical Framework

I test whether older workers are less likely to adopt the computer by estimating a simple linear model for both outcome variables:

$$Y_{it} = \gamma_{a(i)} + \delta X_{it} + \varepsilon_{it}, \tag{D.1}$$

[23] See https://www.ddorn.net/data.htm

The variables of interest are cohort fixed effects γ_a, where a indicates a particular cohort. An observation is a worker i interviewed in year t. I include gender, education, state, occupation, and industry fixed effects interacted with the survey year. Adding education fixed effects accounts for differences in educational attainment across cohorts, which could be a separate channel affecting technology take-up that is not at the core of this exercise. Industry and occupational fixed effects ensure that the regressions do not capture pure sorting.[24]

Note that cohort and age patterns coincide in cross-section, but differ in a panel structure. Focusing on cohort patterns keeps the set of individuals represented by the estimated coefficients constant and, thus, asks "Does it matter how old a subject was when the computer was introduced?" as opposed to "Does the age of a worker matter for current use of a computer?". While the former is focused on the adoption decision, the latter potentially confounds it with life-cycle patterns in technology use.

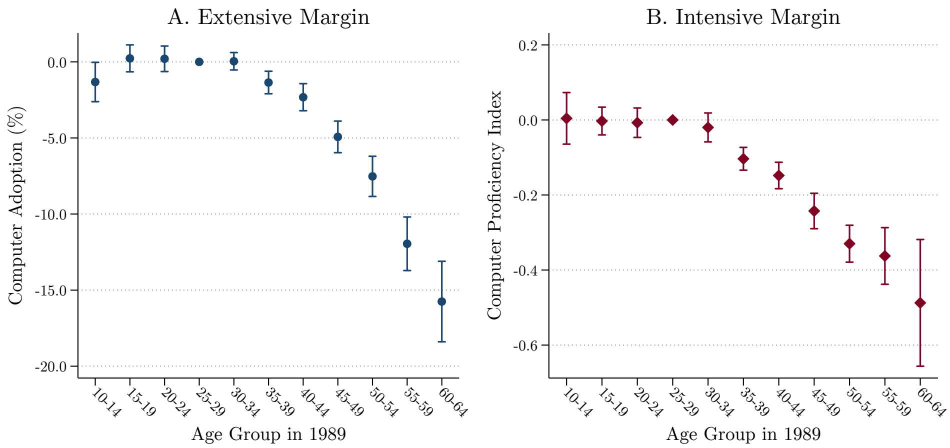
D.3 Results

Panel A of Figure D.1 plots the coefficients for technology adoption, while Tables D.3 and D.3 present the associated regression results. The pattern suggests a monotone decreasing technology adoption rate across cohorts, especially for those aged 40-44 and older in 1989. Panel B confirms a similar pattern for computer proficiency, highlighting that intensive and extensive margin are reinforcing each other. Respondent aged 40-44 in 1989 have a 7.5 percentage points (0.2 tasks) higher computer adoption rate (proficiency index) relative to the cohort age 55-59 in 1989, which constitutes 15% (7%) of the sample mean and 14% (13%) of the sample standard deviation.

In unreported results I confirm cohort patterns as the driving force as opposed to pure life-cycle patterns by simultaneously controlling for age. Furthermore, there does not appear to be any catch-up of older cohorts across survey years, i.e. adoption progresses relatively uniformly across cohorts remaining in the labor market. Finally, note that the CPS does not record employer size or age, which might contribute to the documented patterns if e.g. young firms have a higher technology adoption rate.

[24]Interestingly, the regression tables suggest that sorting appears to be working against the cohort patterns.Older workers tend to work in occupations that use the computer more intensively, flattening the overall cohort profile. This is in line with the evidence provided in Acemoglu and Restrepo (2018), who argue that older workers have a comparative advantage in "white-collar" occupations.

Figure D.1: Older cohorts were slow to adopt the computer



Notes: This figure reports the coefficient estimates for specification (D.1) for computer adoption and proficiency. Regressions include sex, education, industry, occupation, and state fixed effects interacted with survey year. Observations are weighted by CPS Computer and Internet Use Supplement sampling weights. Standard errors are clustered at the industry level.

However, it not necessarily clear that one would want to control for firm age given that the observed sorting of young workers to young firms might be partly driven by (joint) technology adoption decisions (Ouimet and Zarutskie, 2014). Furthermore, the evidence presented focuses on realized patterns, which might differ from "natural" patterns if e.g. employers respond to low technology adoption rates by old workers with more training (Bartel and Sicherman, 1998).

In conclusion, the evidence suggests that older workers adopted the computer at a lower rate in line with the idea that they might be slow to pick up new technologies in general.

Table D.2: Regression Table for Computer Use At Work

	(1)	(2)	(3)	(4)	(5)
			Computer Adoption (%)		
Age 10-14 in 1989	-2.076*	-2.789***	-1.119	-1.169	-1.323**
	(1.229)	(1.065)	(0.711)	(0.712)	(0.655)
Age 15-19 in 1989	0.751	-0.504	0.146	0.130	0.230
	(0.697)	(0.601)	(0.431)	(0.438)	(0.449)
Age 20-24 in 1989	0.457	-0.279	0.080	0.070	0.202
	(0.524)	(0.495)	(0.421)	(0.418)	(0.425)
Age 30-34 in 1989	1.307**	0.907*	0.048	0.014	0.041
	(0.654)	(0.493)	(0.299)	(0.295)	(0.290)
Age 35-39 in 1989	1.165	-0.306	-1.269***	-1.344***	-1.357***
	(1.045)	(0.684)	(0.375)	(0.368)	(0.376)
Age 40-44 in 1989	2.305	0.168	-2.253***	-2.317***	-2.322***
	(1.441)	(0.978)	(0.447)	(0.445)	(0.452)
Age 45-49 in 1989	-1.157	-2.255*	-4.979***	-4.985***	-4.929***
	(1.689)	(1.262)	(0.545)	(0.547)	(0.526)
Age 50-54 in 1989	-4.306**	-4.179***	-7.435***	-7.422***	-7.524***
	(1.807)	(1.464)	(0.646)	(0.647)	(0.670)
Age 55-59 in 1989	-9.081***	-8.938***	-12.237***	-12.138***	-11.953***
	(1.898)	(1.582)	(0.869)	(0.867)	(0.893)
Age 60-64 in 1989	-13.838***	-14.467***	-16.921***	-16.729***	-15.751***
	(2.690)	(2.249)	(1.481)	(1.474)	(1.342)
Gender/Educ. FEs		Yes	Yes	Yes	Yes x Year
Ind./Occ. FEs			Yes	Yes	Yes x Year
State FEs				Yes	Yes x Year
Obs.	207,998	207,998	207,998	207,998	207,983

Note: This table reports the regression coefficients for direct computer use at work. Outcome is an indicator variable taking values 0 and 100 with standard deviation 49.3 and mean 58.34. Age 25-29 in 1989 is the leave out category. Regressions use CPS Computer and Internet Supplement weights and control for year fixed effects. All standard errors clustered at industry level.

Standard Errors in Parenthesis. Significance levels: * 10% , ** 5%, *** 1%.

Table D.3: Regression Table for Tasks Performed With Computer

	(1)	(2)	(3)	(4)	(5)
			Computer Proficiency		
Age 10-14 in 1989	-0.082	-0.089*	-0.010	-0.013	-0.025
	(0.051)	(0.049)	(0.041)	(0.042)	(0.043)
Age 15-19 in 1989	-0.006	-0.035	0.009	0.008	-0.000
	(0.028)	(0.026)	(0.019)	(0.019)	(0.019)
Age 20-24 in 1989	-0.017	-0.039	-0.018	-0.019	-0.021
	(0.025)	(0.024)	(0.023)	(0.024)	(0.024)
Age 30-34 in 1989	-0.031	-0.025	-0.028	-0.030	-0.032
	(0.024)	(0.021)	(0.020)	(0.019)	(0.020)
Age 35-39 in 1989	-0.092***	-0.111***	-0.110***	-0.111***	-0.115***
	(0.026)	(0.021)	(0.017)	(0.017)	(0.018)
Age 40-44 in 1989	-0.104***	-0.135***	-0.157***	-0.159***	-0.164***
	(0.031)	(0.024)	(0.021)	(0.020)	(0.021)
Age 45-49 in 1989	-0.214***	-0.220***	-0.260***	-0.258***	-0.272***
	(0.037)	(0.033)	(0.027)	(0.027)	(0.026)
Age 50-54 in 1989	-0.318***	-0.305***	-0.351***	-0.354***	-0.357***
	(0.045)	(0.039)	(0.027)	(0.027)	(0.027)
Age 55-59 in 1989	-0.317***	-0.318***	-0.372***	-0.371***	-0.385***
	(0.046)	(0.048)	(0.051)	(0.051)	(0.049)
Age 60-64 in 1989	-0.486***	-0.485***	-0.505***	-0.505***	-0.498***
	(0.086)	(0.085)	(0.077)	(0.077)	(0.079)
Gender/Educ. FEs		Yes	Yes	Yes	Yes x Year
Ind./Occ. FEs			Yes	Yes	Yes x Year
State FEs				Yes	Yes x Year
Obs.	109,280	109,280	109,275	109,275	109,160

Note: This table reports the regression coefficients for tasks performed with a computer at work. Outcome is an index variable ranging from 1 to 6 with standard deviation 1.69 and mean 2.8. Age 25-29 in 1989 is the leave out category. Regressions use CPS Computer and Internet Supplement weights and control for year fixed effects. All standard errors clustered at industry level.

Standard Errors in Parenthesis. Significance levels: * 10% , ** 5%, *** 1%.