Aging, Technology Adoption, and Growth

Nils H. Lehr

Boston University *

February, 2021

Abstract

This paper argues that workforce aging has contributed to the productivity slowdown in the US through a technology adoption channel. I document that older workers are slow to adopt new technologies and build an endogenous growth model with overlapping generations and costly learning. The model replicates observed cohort patterns in technology adoption and predicts a scale-back of innovation in response to an aging workforce leading to lower productivity growth in the long run. I confirm the prediction for innovation using a local labor markets approach. Commuting zones with an aging workforce invest less in R&D and produce fewer innovations.

*Email: nilslehr@bu.edu Mailing Address: Boston University, Dept. of Economics, 270 Bay State Rd.,Room B03A, Boston, MA 02215. I benefited from advice from many including Stephen Terry, Pascual Restrepo, David Lagakos, Daniele Paserman, Robert King, Tarek Hassan, Susanto Basu, Adam Guren, and Stefania Garetto as well as the useful comments and suggestions of participants in the Boston University Lunch Seminar, Green Line Macro Meeting, and Lecznar Memorial Lectures.

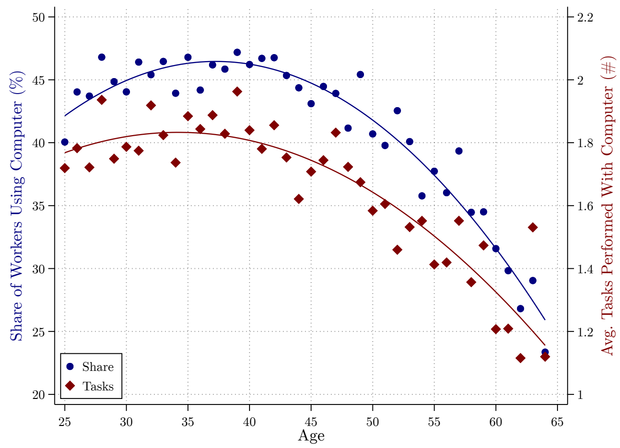
1 Introduction

Productivity growth has been anemic in the US and other developed countries in the previous decade and a half (Gordon, 2016; Syverson, 2017; Andrews et al., 2016). In this paper, I argue that slow productivity growth has been partly driven by an aging labor force. Older workers adopt new technologies at a slower pace and thus, via a composition effect, an aging workforce leads to a lower aggregate technology adoption rate, which in turn implies lower productivity gains in the short-run via adoption lags as well as long-term lower economic growth by reducing the incentives to produce new technologies in the first place.

While demographics have received some attention in the recent macroeconomic literature (Teulings and Baldwin, 2014), few have explicitly considered the age composition of the labor force. Figure 1 highlights that composition might be an important force after all. In particular, there are noteworthy differences in technology adoption across the life-cycle. At the advent of the age of personal computing, young workers were more likely to use a PC at the workplace, and among users, young workers leveraged the PC for a wider range of tasks. For example, around 45% of workers in the first half of their working life (age 25 to 44) adopted the computer at the workplace compared around 35% for those in the second half of their careers. Similarly, among adopters, the former performed around 1.8 tasks with the PC compared 1.5 tasks for the older group.

While this insight might be interesting in and of itself for explaining technology adoption across firms or regions, it is especially relevant to the recent economic performance of developed countries, including the US, due to their rapidly aging workforce. Figure 2 plots the share of age 25-44 subjects among those aged 25-64, a measure I will refer to as Working Young Share (WYS), for the US population, labor force, and employment. In all three cases, the WYS peaked around 1990 followed by a rapid, large scale decline up to 2010 with

Figure 1: Older workers adopted the computer at a slower pace



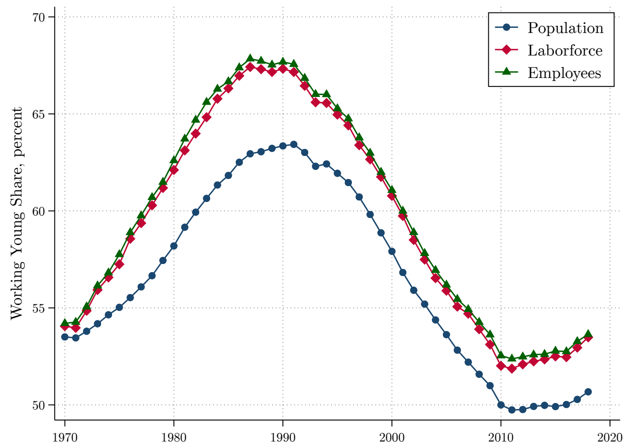
Note: This figure shows the adoption of the computer by workers in 1989. Adoption by workers by age is measured from the CPS October supplement using the share of workers directly using PCs at work and the average tasks performed with a computer when used. The plotted lines show the quadratic fit. See Appendix B.1 for data construction details.

subsequent stabilization. For example, the WYS for the US population peaked at c. 64% in 1990 and declined by more than 10 percentage points to 50% in 2010. These trends are mirrored in many developed countries and UN projections predict them to persist in the medium to long-run (United Nations, 2019).

I proceed in three steps to bolster the claim that workforce aging contributed to the recent productivity slowdown. Firstly, I build on Figure 1 by analyzing computer adoption across age groups in the 1989-2003 period and find strong evidence that older workers were less likely to adopt computers at work. The documented patterns suggest that composition shifts have a meaningful aggregate impact on technology adoption.

Secondly, motivated by this evidence I build an endogenous growth model with overlapping generations and costly technology adoption for workers. The latter naturally gives rise to the observed age patterns in technology adoption due to differences in the remaining time in the labor market. In the model, an

Figure 2: The US workforce has aged rapidly since 1990



Note: This figure shows the WYS for the US population, laborforce 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. See Appendix B.1 for data construction details.

aging workforce results in a lower average adoption rate of new technologies driven by composition effects. Low adoption in turn results in low demand for the associated capital, dampening investment overall, and reducing the profits for new equipment producers. As a result, inventors have lower incentives to develop new technologies in the first place leading to lower aggregate productivity growth.

Comparing the competitive equilibrium to the social planner solution highlights that technology adoption choices amplify the standard static inefficiency arising in endogenous growth models and give them a dynamic flavor. In particular, monopolistic pricing in the competitive equilibrium dampens investment in equipment, which in turn leads to a lower marginal product of labor and, thus, wages. Low wages reduces workers' incentive to learn about the associated technologies leading to a lower aggregate technology adoption rate and, thus, smaller market size for technologies. A policy supporting investment in equipment, thus, has additional static efficiency gains from higher technology adoption rates as well as dynamic gains from increasing the market size and thus profits for new technologies. Finally, the social planner solution highlights that slower productivity growth in response to an aging workforce is optimal, however, optimal growth rates are strictly larger than in the competitive equilibrium.

Finally, I test the model's predictions for R&D investment following a local labor market approach and instrumental variable strategy. Using historical birth rates to instrument for the WYS, I show that commuting zones (CZs) with lower WYS have lower R&D employment and produce fewer patents per capita. I find that a 1 percentage point decline in the local WYS leads to a 0.2 percentage point decline in the R&D employment share and 0.25 lesser citation adjusted patents per 1000 capita. Model and evidence thus jointly suggest that workforce aging has contributed to the observed slowdown in US productivity growth and investment.

This paper contributes to four lines of research. Firstly, I contribute to the growing literature on the recent slowdown in US productivity growth. The existing literature has documented a significant slowdown in productivity growth since at least 2005 together with low investment since around 2000 and explored a range of potential contributing factors.[1] I add to this literature by highlighting the impact of labor force aging through technology adoption as a key factor.

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.[2] I add to this literature by highlighting production-side implications with a focus on technology adoption as driving factor. Technology adoption is

[1] Gordon (2016), Syverson (2017), and Philippon and Gutiérrez (2017) document slow productivity growth and investment. See e.g. Aghion et al. (2019); Liu et al. (2019); Akcigit and Ates (2019); Gordon (2016); Bloom et al. (2019) and Brynjolfsson et al. (2019) for complementary mechanisms explaining these facts.

[2] See the papers in Teulings and Baldwin (2014) and well as Eggertsson et al. (2019). There are two notable exceptions. Firstly, Aksoy et al. (2019) allow for differential research productivity across age groups in their study on the macroeconomic impacts of aging. Secondly, Feyrer (2007) and Maestas et al. (2016) provide evidence that labor force aging is associated with slower productivity growth at the state level.

also a key channel in Acemoglu and Restrepo (2019), who highlight workforce aging as a key contributor to the current wave of automation. I complement their perspective by focusing on worker augmenting technologies and the associated life-cycle pattern of human capital investments.

Thirdly, my paper speaks to the growing literature on firm dynamics and demographics by highlight workforce composition as an important force impacting firm creation. Karahan et al. (2019) and Hopenhayn et al. (2018) argue that the declining labor force growth rate, which is tightly linked to workforce aging, has contributed to declining firm dynamism. I complement their perspective by focusing on the composition of the workforce instead of its size and show empirically that composition has an independent impact on R&D efforts and outputs. In a similar line of inquiry, Engbom (2019) argues that age composition shifts contributed to declining job transition rates, unemployment rates, and entrepreneurship. His mechanism relies on older workers being better matched to their current employment and thus less likely to consider outside options such as entrepreneurship. My framework complements his perspective by adding technology adoption as a key differentiating factor across generations and highlights its separate implications for investment, innovation, and growth. The evidence presented in this paper suggests that this channel is itself an important 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 existing literature documents that individual research and entrepreneurship productivity peaks around age 40-50, which would suggest that workforce aging should have a positive contribution to aggregate entrepreneurship and R&D productivity.[3] In contrast, Derrien et al. (2018) find that local labor markets with a higher share of young workers record higher patenting rates. This paper contributes to the discussion on age and

[3] See Akcigit et al. (2017), Jones (2010), and Jones and Weinberg (2011) for papers on scientific productivity and Azoulay et al. (2020) for entrepreneurship and entrepreneurial success.

innovation by highlighting labor force composition as a driver of new technology demand instead of focusing supply via the inventor or entrepreneur herself.[4]

Section 2 presents direct evidence on age patterns in technology adoption using the computer as a case study. Building on this evidence, Section 3 presents an endogenous growth model with age patterns in technology adoption and derives prediction for investment, innovation, and growth in response to a shift in the age composition towards an older workforce. Section 4 confirms the model's predictions for R&D investment and outputs using a local labor markets approach and Section 5 concludes.

2 Evidence on the Adoption of Computers

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.[5]

[4] Note that this insight is principally orthogonal to the observation that inventor and entrepreneurial success peaks during later ages, however, it implies that the impact of aging on innovation is more complicated than pure composition effects based on inventor or entrepreneurial productivity.

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

2.1 Data

I investigate computer adoption at the workplace using the five CPS Computer and Internet (CIU) Supplement waves between 1989 and 2003. 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. 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.

Besides the CIU specific variables, I use the age and gender of the respondent, state of residency, educational attainment, occupation, and industry. 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. See Appendix B.1 for further details on the data construction and summary statistics.

2.2 Empirical Framework

I substantiate the findings in Figure 1 by estimating a simple linear model for both outcome variables:

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

The core 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 paper. Industry and occupational fixed effects ensure that the regressions do not capture pure sorting.[6] 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.

2.3 Results

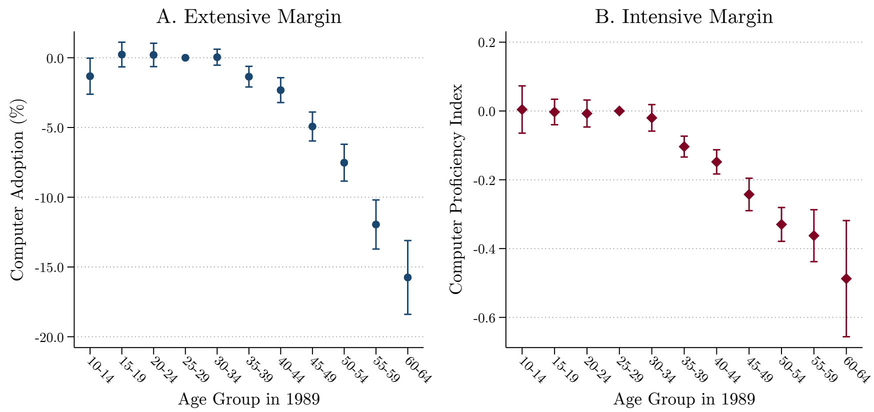
Panel A of Figure 3 plots the coefficients for technology adoption.[7] The pattern suggests a monotone decreasing technology adoption rate across cohorts, especially for those aged 40-44 and older in 1989. Panel B confirms a similar pattern for computer proficiency, highlighting that intensive and extensive 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.

I report robustness controlling for age groups and confirm cohort patterns as the driving force as opposed to pure life-cycle patterns in Appendix C.1. 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. This confirms that the evidence in Figure 3 does not reflect pure timing differences across cohorts.

[6] As shown in the regression tables in the appendix, sorting appears to be working against the cohort patterns. In other words, 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 (2019), who argue that older workers have a comparative advantage in “white-collar” occupations.

[7] See Appendix C.1 for regression tables and robustness checks.

Figure 3: Older cohorts were slow to adopt the computer



Notes: This figure reports the coefficient estimates for specification (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. See Appendix B.1 for data construction details.

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. 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, 2002).

In conclusion, the evidence suggests that older workers adopt new technologies at a lower rate. This finding is at the heart of the model developed in the next section.

3 A Model of Aging and Technology Adoption

Motivated by the evidence in the previous section, I develop an endogenous growth model featuring cohort effects and discuss the impact of demographic change on innovation and investment. The model builds on the standard expanding varieties growth model as in Romer (1990) and extends it in two directions.[8] Firstly, I introduce demographics using a standard overlapping generations structure, and, secondly, technology adoption is made an explicit choice on part of workers. The resulting model delivers strong, testable predictions on the impact of population aging on output, investment, and R&D efforts.

3.1 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 as the sole input. 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.

Technology takes a central position in this paper, so I first discuss the related notation at length. I will denote the set of technologies available and the set of new inventions at time t as A_t and a_t respectively. The stock of aggregate technologies evolves cumulatively according to

$$A_t = a_t + A_{t-1} \tag{2}$$

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

In words, the stock of technologies available is simply the sum of technologies at time $t-1$ and new inventions in the current period.

3.1.1 Households

There is a representative household maximizing

$$\sum_{s=0}^{\infty} \beta^s (1+n)^s \ln(c_{t+s}), \tag{3}$$

where β is the time discount factor, n is the population growth rate, and c_t is per capita consumption.[9]

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

$$(1+n)b_{t+1} = (1+r_t)b_t + w_t - h_t - c_t. \tag{4}$$

For simplicity, I will restrict savings to be non-negative, $b_{t+1} \geq 0$, instead of a more rigorous no-Ponzi condition. This constraint will not be binding in the equilibria considered below.

The household itself is composed of two active 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:

[9] 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.

$$s_y = \frac{1+n}{2+n-p} \tag{5}$$

My analysis below will focus on comparative statics across Balanced Growth Path equilibria with different population growth rate n and thus abstracts from transition dynamics induced by time-varying birth rates. Note that comparative statics for n are the appropriate analysis when considering the US. As discussed in Appendix Engbom (2019) and Karahan et al. (2019), the demographic patterns in Figure 2 are primarily driven by declining fertility rates.

Having introduced the demographic structure, we can now discuss technology adoption, which is modeled as a costly, one-off investment on part of the household. In particular, 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 technologies in their skill set, so a larger skill set translates into a larger effective labor supply.[10]

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:

$$\ell_t(a) = s_y \ell_{yt}(a) + (1-s_y)\ell_{ot}(a). \tag{6}$$

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

$$w_t = s_y \int_{A_t} \ell_{yt}(a) W_t(a) da + (1-s_y) \int_{A_t} \ell_{ot}(a) W_t(a) da \tag{7}$$

Knowledge does not come for free, however. All technologies are subject to

[10] As noted below, this is a simple extension of the Romer (1990) framework, which implicitly assumes that workers can work with all technologies at once. I extend this framework by making technologies adoption an explicit choice.

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.[11]

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

$$h_{yt} = A_t \int_0^{n_{yt}} n dF(n) \quad \text{and} \quad \ell_{yt}(a) = F(n_{yt}). \tag{8}$$

The formulation takes advantage of homogeneous adoption costs, which guarantee that the share of adopters is identical across available technologies.[12]

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

[11] 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.

[12] 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.

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}:

$$h_{ot} = A_{t-1} \int_0^{n_{ot}} \mathbb{1}\{n_{yt-1} < n\} n dF(n) + a_t \int_0^{n_{ot}} n dF(n). \tag{9}$$

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.

$$\ell_{ot}(a) = \begin{cases} F(\max\{n_{yt-1}, n_{ot}\}) & \text{if} \quad a \in A_{t-1} \\ F(n_{ot}) & \text{if} \quad a \in a_t. \end{cases} \tag{10}$$

Total technology adoption costs are the aggregate across generations:

$$h_t = s_y h_{yt} + (1 - s_y) h_{ot}. \tag{11}$$

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.

3.1.2 Final Good Producer

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.[13]

[13] 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

$$y_t = \int_{A_t} \ell_t(a)^{1-\alpha} k_t(a)^{\alpha} da. \tag{12}$$

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

$$\max \quad y_t - \int_{A_t} W_t(a)\ell_t(a)da - \int_{A_t} P_t(a)k_t(a)da \qquad \text{s.t.} \quad (12). \tag{13}$$

3.1.3 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

$$\max P_t(a)k_t(a) - \psi k_t(a), \quad \text{s.t.} \quad P_t(a) = \alpha \left(\frac{\ell_t(a)}{k_t(a)}\right)^{1-\alpha}. \tag{14}$$

production function simplifies to

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

which is the standard form once we undo the normalization by population size.

3.1.4 Innovation Sector

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 a_{t+1} according to the simple linear production function:[14]

$$a_{t+1} = \varphi_0 x_t. \tag{15}$$

To simplify the exposition, I will directly assume that the innovation sector is governed by two equations. Firstly, equation (16) 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:[15]

$$\varphi_0 v^0_{t+1} = \left(\frac{1+r_{t+1}}{1+n}\right), \tag{16}$$

where v^0_{t+1} is the expected net present value of profits from a new invention and φ_0 the research productivity.[16] Appendix A.2 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

$$r_t b_t = \int_{A_t} \pi_t(a) da. \tag{17}$$

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

[14] Formulating the production function in per capita terms neutralizes strong market size effects from population growth (see e.g. Jones (1995a,b)). This simplifies the exposition greatly and allows me to focus on balanced growth path differences. The main results below will still be in effect in a semi-endogenous growth setup as in e.g. Jones (1995b), 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.

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

[16] See Appendix A.2 for a full definition.

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.

3.1.5 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.

$$y_t = \int_{A_t} \psi k_t(a) da + h_t + x_t + c_t. \tag{18}$$

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

$$x_t = (1+n)b_{t+1} - b_t. \tag{19}$$

3.1.6 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*

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

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.*

3.2 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 A.2.

Lemma 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 $\frac{1+r}{1+n} > 1$.

Proof. All proofs are deferred to Appendix A. □

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

$$n_y = \mathscr{W}\left(1+\frac{1-p}{1+r}\right) \quad \text{and} \quad n_o = \mathscr{W}. \tag{20}$$

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 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.

3.2.2 Firm Profits and the Value of Innovation

Having solved the worker problem, we can next turn our attention to the intermediary problem.

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

$$\pi_t(a) = \tilde{\pi}\ell_t(a). \tag{21}$$

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

$$v^0 = \tilde{\pi}\left(\ell^N + \left(\frac{1+n}{r-n}\right)\ell^E\right), \tag{22}$$

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.

3.2.3 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 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 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 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 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 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.*

3.2.4 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 A and focus on its implications here:

Proposition 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 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 6 is the social planner equivalent to Proposition 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.

4 Evidence from Local Labor Markets

I test the model's core prediction regarding innovation and aging following the local labor market approach pioneered by Autor and Dorn (2013). In particular, I construct measures of innovation inputs and outputs at the CZ level and investigate their response to changing demographics during the 1980 to 2010 period. CZs are a partition of counties based on commuting patterns across county borders and are developed to capture the relevant labor market for workers within a CZ (Tolbert and Sizer, 1996). I follow Autor and Dorn (2013) in using the 1990 delineation of CZs.

4.1 Data

The empirical exercise combines three separate sources of information on local labor markets. Firstly, I use the decennial Census/ACS for the 1980-2010 period to construct a range of variables describing the local labor force composition and R&D employment. Secondly, I use the US PTO Patentsview bulk files to characterize patenting and patent citations for local inventors. Finally, I construct an instrument for population composition across local labor markets based on historical fertility rates using data from historic censuses, NBER Vitality statistics, and NBER CDC SEER data. See Appendix B for further details.

I use two measures for innovative activity. Firstly, I proxy for R&D inputs using the share of workers employed in R&D occupations, which I broadly define as scientists and engineers.[17] I restrict my sample to individuals aged 25 to 64 and, following Acemoglu and Autor (2011), focus on Full-Time Full-Year (FTFY) workers only. Secondly, I proxy for R&D outputs using citation-adjusted patent grants to local investors. As is standard in the literature, I restrict citations to the first five years since the patent was granted and record patents in their application year.[18] I normalize citation-adjusted patent counts by population to make them comparable across CZs. The final variable records patents per 1000 capita, but I will refer to it as per capita.

My key summary statistic for workforce aging is the Working Young Share (WYS), which I define as the share of people aged 25 to 44 within the overall population of age 25 to 64:

$$\text{WYS}_{CZ,t} = \frac{\text{Population Age 25-44}_{CZ,t}}{\text{Population Age 25-64}_{CZ,t}}.$$

The classification into young and old is motivated by the empirical evidence in Section 2, which highlighted a fast drop-off in adoption rates after age 44, and conveniently splits the working-age population in half.

4.2 Empirical Framework

The core prediction of the model developed in the previous section is that an increase in the WYS should lead to an increase in innovation activity. For both outcome variables, R&D employment and patents per capita, I estimate

[17] Employment is an important driver of overall R&D costs: According to the NSF's Business R&D and Innovation Survey, labor had a cost-share of 66.9% for domestic R&D in US companies in 2015 (Business R&D and Innovation Survey, 2015). While it would have been preferable to have direct measures of R&D expenditures, these are not (publicly) available for detailed geographic units.

[18] Citation-weighting is standard in the innovation literature on patenting and is meant to capture the quality of the innovation (Pakes, 1985; Harhoff et al., 1999; Hall et al., 2005; Griliches, 1990). More recently, Kogan et al. (2017) confirm that citations are positively correlated with the market valuation in a large set of corporate patents. I confirm that my results are not sensitive to the citation weighting by using raw patent counts instead as a robustness check.

(23) to confirm this prediction:

$$\Delta Y_{CZ,t} = \alpha_{CZ} + \gamma_t + \beta \Delta \text{WYS}_{CZ,t} + \varepsilon_{CZ,t}, \tag{23}$$

where $\Delta Y_{CZ,t}$ is the difference between t and $t-1$ for variable $Y_{CZ,t}$ in a given CZ. Running the model in changes safeguards against preexisting level differences. The full specification includes CZ and year fixed effects, which control for differential trends and common movements across years respectively. In specifications without CZ fixed effects I control for a range of initial condition measured in 1980 to ensure that I am not capturing long-run differences.[19] I follow Autor and Dorn (2013) and weight observations by population size. I use the 1980 working age population to ensure that commuting zones have a constant weight.

While running the specification in changes with a rich set of fixed effects alleviates some of the identification issues, there remain at least two important concerns with the OLS regression.

Firstly, the sorting of young (workers) across places and employment opportunities is likely highly endogenous. Young workers are more mobile, both geographically and across employment opportunities, and tend to sort into nascent industries and young firms.[20] This can lead to an upwards or downwards bias in the regression above. For example, if young workers move towards booming places, we might expect a positive bias in the regression. On the other hand, if we believe that young workers predominantly sort into CZ with many young, financially constraint firms, then the coefficient might be downwards biased.

Secondly, the local labor market is likely to be an imprecise proxy for the relevant workforce for R&D decisions. For example, R&D opportunities are likely

[19]These include the initial WYS, the share of women, non-white, and working-age population with a college degree, size of the working-age population (in logs), and region fixed effects. Note that all of these drop out once I include CZ fixed effects.

[20]See (Kaplan and Schulhofer-Wohl, 2017) for geographic mobility, and Shimer (2001); Ouimet and Zarutskie (2014), and Engbom (2019) for sorting and job switching.

to be unequally distributed across industries such that the relevant workforce is restricted to the workers in the particular industry. Furthermore, R&D might be driven more by the predicted workforce than the actual workforce. In other words, the research specification in the model crucially depends on firms being able to predict workforce changes. In contrast, firms might not be as sensitive to realized population demographics as they might not necessarily capture the relevant workers available once a technology is developed.
I address both of these concerns following an instrumental variable strategy.

4.3 Instrumental Variable Strategy

I instrument for the WYS using county-level births from 1920 to 1990 following an expanding literature on demographics in macroeconomics.[21] In particular, to calculate the artificial WYS based on births I aggregate births 25 to 44 years ago and divide them by birth 25 to 64 years ago in the same commuting zone. I adjust birth counts using 1980 mortality rates across age groups in the aggregation. The instrument captures population dynamics in a world with constant mortality rates and without geographic mobility.

$$\text{WYS}_{CZ,t}^{Instr.} = \frac{\sum_{\tau=25}^{44} \text{Births}_{CZ,t-\tau} \times \text{Survival rate}(\tau)}{\sum_{\tau=25}^{64} \text{Births}_{CZ,t-\tau} \times \text{Survival rate}(\tau)}.$$

Note that the instrument relies on data that is realized at least 25 years before the actual observation. Furthermore, the regression specification is in differences with time and CZ fixed effects, addressing e.g. concerns about permanent level or trend differences in fertility rates across CZs.
The exclusion restriction can be summarized as follows:

> *Conditional on CZ and time fixed effects, changes in the population composition by birth are only related to the outcome variable*

[21] See e.g. Shimer (2001), Acemoglu and Restrepo (2019), Derrien et al. (2018), Engbom (2019), and Karahan et al. (2019)

of interest through their impact on changes in the working young share.

A natural concern for this kind of instrument is the question of where the remaining variation is derived from and whether it still potentially violates the exclusion restrictions. I will address this concern for both outcome measures by showing that coefficients remain stable even when controlling for a range of variables that might be associated with changing birth rates (Angrist and Pischke, 2009). See Appendix C.2.1 for further discussion.

4.4 Did workforce aging reduce innovation in the US?

Table 1 reports the results for the R&D employment share. The OLS coefficient in the full model in column (5) is positive and highly significant, indicating a robust correlation between changes in the WYS and R&D employment. The IV coefficient in column (6) is about twice the size and remains highly significant. A one standard deviation higher WYS leads to a 1 standard deviations higher R&D employment. This is not driven by pure composition effects: Columns (7) and (8) report additional results for an age-adjusted R&D employment share. The coefficient estimates remain stable, highlighting that the estimated effects are driven by within age group changes.

The IV results for patenting confirm a strong, significant effect of the WYS on patenting. The coefficients for the full specification, reported column (6) of Table 2, indicate that a one standard deviation increase in the WYS leads to 1.1 standard deviations higher citation adjusted patent grants per capita. In contrast, I do not find strong evidence for a positive association with the WYS in the OLS specification. I report robustness checks using alternative patenting measures in Appendix Table C.6 and confirm a robust relationship between patenting and the WYS in the IV specification.

For robustness, I report results for alternative commuting zone weights in Appendix Table C.8 and confirm very stable results for R&D employment and patenting using no weights or simple working age population weights instead. Furthermore, Appendix Table C.9 reports results allowing for time-varying regional and state trend. The results confirm a strong positive relationship, however, the instrument becomes unreliable when allowing for time-varying state-specific trends.

The regression results raise the question of why the IV coefficients are so different from the OLS results. One potential explanation is the that instrument captures a long-run, highly predictable path of demographics that is key to firm decision making, while the observed WYS has a strong noise component.[22] I attempt to address some of the remaining concerns below by showing that my results are not driven by correlated trends in e.g. education or population growth.

4.5 Alternative Mechanisms

A natural concern when using an instrumental variable strategy are violations of the exclusion restriction. In particular, one might be concerned that the instrument could be correlated with other underlying trends that are the "true" causal driver of my empirical results. Importantly, the instrument is mechanically related to the labor force growth rate, which has been linked to innovation and entrepreneurship in other contexts (Jones, 1995a; Karahan et al., 2019; Hopenhayn et al., 2018). Furthermore, one might be concerned that birth rates are linked to female educational attainment and racial composition, which could independently affect innovation (Bailey, 2006; The Center for Disease Control and Prevention, 2018).

[22] See alsoShimer (2001) and Engbom (2019), who find similar differences for the instrument.

Table 1: Local R&D Employment: Main Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Second stage	**Δ R&D emp.**							
Δ WYS	0.051***	0.037***	0.109***	0.159***	0.131***	0.207***	0.118***	0.251***
	(0.008)	(0.010)	(0.024)	(0.057)	(0.025)	(0.049)	(0.028)	(0.058)
First stage	**Δ WYS**							
Δ WYS (instr.)		1.062***		0.205***		0.263***		0.263***
		(0.040)		(0.042)		(0.038)		(0.038)
Age-adjusted							Yes	Yes
Initial conditions	Yes	Yes	Yes	Yes				
Fixed	Region	Region	Region	Region	CZ	CZ	CZ	CZ
Effects			Year	Year	Year	Year	Year	Year
F-statistic (1st)		700.8		24.0		49.0		49.0
Observations	2,166	2,166	2,166	2,166	2,166	2,166	2,166	2,166

Note: This table reports the OLS and IV coefficient estimates for specification (23) for R&D employments among Full-Time Full-Year employees. Odd columns present OLS results, while even columns present the coefficients for the IV specification. The top panel reports the second stage or OLS results and the bottom panel first stage results. Columns(1)-(4) control for initial conditions in 1980 including the share of working young, women, non-white, and college graduates as well as the working-age population. Columns (7) and (8) present results for R&D employment shares adjusted to hold the age composition constant at 1980s levels. See Appendix B for data construction details. The reported F-statistic is the Sanderson-Windmeijer multivariate F-test of excluded instruments. Observations are weighted by 1980 working-age population and standard errors clustered at the state level.

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

Table 2: Local Patenting: Main Results

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
Second stage			**Δ Patents p.c.**			
Δ WYS	0.065***	0.053***	0.003	0.220**	0.024	0.254**
	(0.007)	(0.009)	(0.018)	(0.093)	(0.021)	(0.100)
First stage			**Δ WYS**			
Δ WYS (instr.)		1.064***		0.208***		0.267***
		(0.040)		(0.042)		(0.038)
Initial conditions	Yes	Yes	Yes	Yes		
Fixed	Region	Region	Region	Region	CZ	CZ
Effects			Year	Year	Year	Year
F-statistic (1st)		721.3		24.5		49.2
Observations	1,924	1,924	1,924	1,924	1,924	1,924

Note: This table reports the OLS and IV coefficient estimates for specification (23) for citation-weighted patent grants per 1000 capita. Odd columns present OLS results, while even columns present the coefficients for the IV specification. The top panel reports the second stage or OLS results and the bottom panel first stage results. Columns(1)-(4) control for initial conditions in 1980 including the share of working young, women, non-white, and college graduates as well as the working-age population. See Appendix B for data construction details. The reported F-statistic is the Sanderson-Windmeijer multivariate F-test of excluded instruments. Observations are weighted by 1980 working-age population and standard errors clustered at the state level.

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

I try to alleviate some of these concerns following the bad control approach suggested in Angrist and Pischke (2009). In particular, I estimate

$$\Delta Y_{CZ,t} = \alpha_{CZ} + \gamma_t + \beta \Delta \widehat{\text{WYS}}_{CZ,t} + \delta X_{CZ,t} + \varepsilon_{CZ,t}, \tag{24}$$

where $\Delta \widehat{\text{WYS}}_{CZ,t}$ is the predicted change in WYS based on the first stage to specification (23) and $X_{CZ,t}$ includes measures of population growth and changes in educational, gender, and racial composition. Appendix Table C.5 confirms that the estimated β for both innovation measures is effectively un-changed by the inclusion of these additional control variables. Thus, these alternative mechanism do not appear to be a driving force behind my baseline results.

4.6 Interpretation

A second point of contention might be the interpretation of the results. In particular, note that regression equation (23) itself is inconclusive about the channel through which the WYS affects R&D inputs and outputs. While my model emphasizes a demand channel, a supply-side interpretation cannot be ruled out ex-ante. In particular, Derrien et al. (2018) document patterns for patenting in line with the results above and argue in favor of a supply channel.[23] Note, however, that a supply-side channel is somewhat at odds with the existing large literature on individual performance of researchers and entrepreneurs along the life-cycle, which generally finds that the individual productivity as researcher and entrepreneur peaks around age 40 to 50 (Jones, 2010; Jones and Weinberg, 2011; Akcigit et al., 2017; Azoulay et al., 2020).

I try to address this concern in two ways. Firstly, I construct an alternative measure of R&D employment that holds constant the local age composition age at 1980's levels and relies on within age changes in the propensity to become an R&D worker only. This addresses any mechanical effect coming from changing age composition and, thus, should address any inherent differences in research productivity. Columns (7) and (8) in Table 1 report the associated results. In contrast to a pure supply based theory, which would predict significantly lower coefficients for the adjusted measure, I find marginally larger effects.

Secondly, I report additional results for R&D employment as well as the associated wages in Appendix Tables C.7. The results confirm that the estimated

[23] Note that while their framework is very similar to the one proposed above, there are important differences. Firstly, they restrict their attention to reduced form evidence only using an instrument similar to the one proposed below, while my results rely on a full 2SLS framework. Secondly, they only consider R&D outputs, while I consider inputs as well, and, finally, they restrict their attention to specification with either time fixed effects or CZ fixed effects, while I can to use both.

effects hold when focusing on private sector employment only and when excluding employees in the educational sector. Thus, my results are truly driven by a private sector response and not by demand for educators or similar channels. Finally, I report results for tradable industries only, whose demand is likely to be relatively independent of the local labor force composition. My results confirm that there does not appear to be any significant effect on R&D employment in tradable industries, which supports a demand driven interpretation and is somewhat at odds with a supply driven explanation as local skill supply should be independent of the industry in question.

Results for wages mirror the result for employment overall. This further support an interpretation of increasing demand for R&D employees in a young commuting zone. Furthermore, it rules out a price-effect interpretation as that young R&D workers might be cheaper leading to higher employment. In contrast, we observe higher employment as well as higher wages.

4.7 Magnitudes

Finally, I want to briefly discuss the magnitudes implied by the estimated coefficient. This discussion should naturally be prefaced by mentioning that the estimated elasticities are partial equilibrium in nature and, thus, cannot be simply aggregated to receive meaningful aggregate effect sizes (Nakamura and Steinsson, 2018). The results suggest that a one standard deviation increase in the WYS leads to a 1.4 percentage points higher R&D employment share and 0.5 more patents per 1000 capita. Note that these estimates are quite large. The respective sample means for both variables in 2010 was 2.1% and 0.55.

The large effect size naturally raises the question as to the general equilibrium forces dampening the partial equilibrium effects. A natural candidate for this is reallocation. Part of the estimates response might be R&D facilities getting reallocated to CZs with young workers, which would naturally lead to a large partial equilibrium response even if the general equilibrium effects are significantly smaller. Furthermore, responses at the margin might overestimate the impact of large changes, e.g. because of decreasing returns to scale in R&D or adjusting factor prices.

5 Conclusion

The US has experienced a large demographic shift since 1990 that is forecast to continue. While 64% of the working-age population were below age 45 in 1990, only 51% were in 2018 and only 52.5% are projected to be in the medium term. This paper argues that workforce aging has direct implications for investments in new technologies and innovation. Empirically, I document that older workers were slow to pick up the computer at the workplace and that commuting zones with aging workforce invest less in R&D and produce fewer patents per capita. I show that this is qualitatively in line with a simple endogenous growth model with overlapping generations in which workers have technology adoption costs. In the model, older workers adopt fewer new technologies as they are closer to retirement, and thus have less to gain from them over their remaining time in the labor market. An aging society thus has a lower technology adoption rate, and thus lesser demand for new technologies. The innovation sector reacts to this decline in demand by producing fewer new technologies leading to a slowdown in aggregate productivity growth.

A natural question is whether we should be concerned about this. The model gives us some clues, which echo the conclusion in Vollrath (2020): Slow growth is not necessarily bad. In particular, the social planner solution also features a slow down in innovation and productivity growth in face of lower population growth and aging. The reason for this is simple: Fewer young people implies that future resources are less valuable to a utilitarian social planner.

References

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

_ **and Pascual Restrepo**, "Demographics and Robots," 2019.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li, "A Theory of Falling Growth and Rising Rents," *Federal Reserve Bank of San Francisco, Working Paper Series*, 2019, (August), 01–43.

Akcigit, Ufuk and Sina Ates, "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," *NBER Working Paper*, 2019, *25755.*

_ **, John Grigsby, and Tom Nicholas**, "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age," *National Bureau of Economic Research*, 2017.

Aksoy, Yunus, Henrique S Basso, Ron P Smith, and Tobias Grasl, "Demographic structure and macroeconomic trends," *American Economic Journal: Macroeconomics*, 2019, *11* (1), 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," Technical Report 5 dec 2016.

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* (509), 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* (5), 1553–1597.

_ , Frank Levy, and Richard J. Murnane, "The skill content of recent technological change: An empirical exploration," *Quarterly Journal of Economics*, nov 2003, *118* (4), 1279–1333.

_ , Lawrence F. Katz, and Alan B. Krueger, "Computing inequality: Have computers changed the labor market?," *Quarterly Journal of Economics*, 1998, *113* (4), 1169–1213.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda, "Age and High-Growth Entrepreneurship," *American Economics Review: Insights*, 2020, *2* (1), 65–82.

Bailey, Martha J, "More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply," *The Quarterly Journal of Economics*, feb 2006, *121* (1), 289–320.

Bartel, Ann P. and Nachum Sicherman, "Technological Change and the Skill Acquisition of Young Workers," *Journal of Labor Economics*, 2002, *16* (4), 718–755.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen, "What Drives Differences in Management Practices?," *American Economic Review*, may 2019, *109* (5), 1648–1683.

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* (1), 339–376.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics," in "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* (1), 137–198.

Derrien, Francois, Ambrus Kecskes, and Phuong-Anh Nguyen, "Labor Force Demographics and Corporate Innovation," 2018.

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), 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* (1), 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* (3), 511–529.

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

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

Griliches, Zvi, "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 1990, *28* (4), 1661–1707.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg, "Market Value and Patent Citations," *The RAND Journal of Economics*, 2005, *36* (1), 16–38.

Harhoff, Dietmar, Francis Narin, F. M. Scherer, and Katrin Vopel, "Citation Frequency and the Value of Patented Inventions," *The Review of Economics and Statistics*, 1999, *81* (3), 511–515.

Hopenhayn, Hugo A., Julian Neira, and Rish Singhania, "From Population Growth To Firm Demographics :," *NBER Working Papers*, 2018.

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

_ **and Bruce A. Weinberg**, "Age dynamics in scientific creativity," *PNAS*, 2011, *108* (47), 18910–18914.

Jones, Charles, "R & D-Based Models of Economic Growth," *Journal of Political Economy*, 1995, *103* (4), 759–784.

_ , "Time Series Tests of Endogenous Growth Models," *The Quarterly Journal of Economics*, 1995, *110* (2), 495–525.

Kaplan, Greg and Sam Schulhofer-Wohl, "Understanding the Long-Run Decline in Interstate Migration," *International Economic Review*, 2017, *58* (1), 57–94.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin, "Demographic Origins of the Startup Deficit," 2019.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, “Technological Innovation, Resource Allocation, and Growth,” *The Quarterly Journal of Economics*, may 2017, *132* (2), 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* (2), 797–849.

Liu, Ernest, Atif Mian, and Amir Sufi, “Low Interest Rates, Market Power, and Productivity Growth,” 2019.

Maestas, Nicole, Kathleen J. Mullen, and David Powell, “The Effect of Population Aging on Economic Growth, The Labor Force And Productivity,” *NBER Working Paper*, 2016, *22452.*

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

Nakamura, Emi and Jón Steinsson, “Identification in Macroeconomics,” *Journal of Economic Perspectives*, aug 2018, *32* (3), 59–86.

Ouimet, Paige and Rebecca Zarutskie, “Who works for startups? The relation between firm age, employee age, and growth,” *Journal of Financial Economics*, 2014, *112* (3), 386–407.

Pakes, Ariel, “On Patents, R&D, and the Stock Market Rate of Return,” *Journal of Political Economy*, 1985, *93* (2), 390–409.

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* (5).

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

Schaller, Jessamyn, Price Fishback, and Kelli Marquardt, "Local Economic Conditions and Fertility from the Great Depression through the Great Recession †," *American Economic Review: Papers & Proceedings*, 2020, pp. 236–240.

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*, aug 2001, *116* (3), 969–1007.

Syverson, Chad, "Challenges to Mismeasurement Explanations for the US Productivity Slowdown," *Journal of Economic Perspectives*, may 2017, *31* (2), 165–186.

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

The Center for Disease Control and Prevention, "Births: Final Data for 2017," *National Vital Statics Reports*, 2018, *67* (8).

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

United Nations, "2019 Revision of tWorld Population Prospects," Technical Report 2019.

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

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

Online Appendix

A Model Appendix

A.1 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.

A.2 Competitive Equilibrium

A.2.1 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 $\mathscr{K}$ and equilibrium task wage $\mathscr{W}$ via the first order conditions of the final good producer:

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

Plugging in the definition of $\mathscr{K}$ and $\mathscr{P}$ yields the expression in Lemma 1 for the task wage.[24] Note that we can already solve for firm profits nd the value of a new invention conditional on household adoption:

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

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

[24]Note that $\partial\mathscr{W}/\partial\mathscr{P} < 0$, 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.

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

Lemma (Restatement of Lemma 1)**.** The BGP interest rate satisfies $1+r = \frac{1+g}{\beta}$. As long as $g \geq 0$, the effective discount rate of the economy satisfies $\frac{1+r}{1+n} > 1$.

Proof. 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. □

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:

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

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

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

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

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

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 doesn't 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.

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

$$n_y = \mathscr{W}\left(1+\frac{1-p}{1+r}\right) \quad \text{and} \quad n_o = \mathscr{W}.$$

Proof. See derivations above. □

Corollary (Restatement of Corollary 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.

Proof of Corollary 1. See derivations above for part (a).
For part (b) note that it follows immediately from (20) 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 $\mathscr{A}^y$ and $\mathscr{A}^o$ respectively, is given by:

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

Given that $n_o < n_y$, it follows immediately that $\mathscr{A}_y > \mathscr{A}_o$ 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

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

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:

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

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 $\mathscr{W} - n_o = 0$ by definition of the adoption threshold. Thus $\mathscr{W} - n$ 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

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

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 $F(n_y)$. □

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

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

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

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

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

Lemma 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. 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 remains to be shown then that $\tilde{c} > 0$ on the balanced growth path. Note that I've dropped time indices due to the focus on the balanced growth path where these quantities are constant. For this consider first the net resources by intermediary firms and innovation sector. For this note that by the research arbitrage equation we have

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

Using the Euler equation, we can simplify this term further to

$$\tilde{x} = \frac{g}{\varphi_0} = \frac{g}{1+g} \beta (1+n) \alpha \mathscr{W} \left(s_y F(n_y) + (1-s_y) F(n_o) + \frac{\beta(1+n)}{1+g-\beta(1+n)} F(n_y) \right)$$

Let π^N be the profits of a new firm and pi^E the profits of an old firm, then we can express this as

$$\tilde{x} = \frac{g}{1+g} \beta (1+n) \left(\pi^N + \frac{\beta(1+n)}{1+g-\beta(1+n)} \pi^E \right)$$

In turn, current profits can be expressed as

$$\tilde{\pi} = \frac{1}{1+g} \pi^E + \frac{g}{1+g} \pi^N$$

Using that $(1+n)\beta < 1$, it can be shown that $\tilde{\pi} > \tilde{x}$. In particular, it is straight-forward to show that

$$\frac{g}{1+g} \pi^N > \frac{g}{1+g} \beta (1+n) \pi^N.$$

Furthermore, the inequality

$$\frac{1}{1+g}\pi^E > \frac{g}{1+g}\frac{(\beta(1+n))^2}{1+g-\beta(1+n)}\pi^E$$

can be rearranged to

$$g + (1-\beta(1+n)) > g(\beta(1+n))^2,$$

which always holds true due to $\beta(1+n) < 1$. Together, both inequalities imply $\tilde{\pi} - \tilde{x} > 0$. Finally, we want to show that $\tilde{w} - \tilde{h} \geq 0$ to guarantee $\tilde{c} > 0$.

I will break this question down into four parts. In particular, one can break down the overall term into

$$\begin{aligned}
A &= s_y \int_0^{n_o} (\mathscr{W} - n) dF(n) \\
B &= s_y \int_{n_o}^{n_y} (\mathscr{W} - n) dF(n) \\
C &= (1-s_y)\frac{1}{1+g}\mathscr{W} F(n_y) \\
D &= (1-s_y)\frac{g}{1+g}\int_{n_o}^{n_y} (\mathscr{W} - n) dF(n)
\end{aligned}$$

A and B concern the labor income of the young, C and D that of the old. Note that A and D are (weakly) positive by definition of n_o. Thus, it remains to be shown that $C + B > 0$. Firstly note

$$B \geq -s_y \frac{1-p}{1+r}\mathscr{W}(F(n_y) - F(n_o)) = -(1-s_y)\frac{\beta(1+n)}{1+g}\mathscr{W}(F(n_y) - F(n_o))$$

Note that the inequality follows from $n \leq n_y$ in the interval concerned. The equality then follows by plugging in the Euler equation and using the relative size of cohorts.
It then immediately follows that

$$C+B \geq \frac{(1-s_y)}{1+g}\mathscr{W}\left(F(n_y)-\beta(1+n)(F(n_y)-F(n_o))\right)>0,$$

where the first inequality follow by plugging in C and using the inequality for B established above. The second inequality then follows from $(1+n)\beta<1$ and $F(n_o)\geq 0$. Together with $A,D\geq 0$, this implies $\tilde{w}-\tilde{h}>0$. Thus, $\tilde{c}>0$.
Finally, it is straight-forward to 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

$$\tilde{\lambda}_{t+s}=\tilde{\lambda}_t\left(\frac{1+n}{1+r}\right)^s=\tilde{\lambda}_t\left(\frac{\beta(1+n)}{1+g}\right)^s \tag{25}$$

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.

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

Proof. Firstly, note that Lemma 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. □

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

Proof. 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. □

Proposition (Restatement of Proposition 3)**.** *Holding constant the population growth rate, an exogenous decline in the working young share decreases the average 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.*

Proof. I will start the proof from the last point. Consider the research arbitrage equation:

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

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

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

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, $a_t \ell^N \mathscr{K}$, is increasing in s_y as well. □

Proposition (Restatement of Proposition 4)**.** *An decrease in the population growth rate, which mechanically leads to a decrease in the working young share, decreases the average adoption rate for new and overall technologies, investment into new technologies relative to old technologies, increases the value of new inventions, the interest rate, and the economy's productivity growth rate.*

Proof. 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. □

A.3 Social Planner Solution

A.3.1 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*

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

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.*

A.3.2 Derivations and Proofs

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 5. On a social planner BGP, the social planner chooses a higher capital-labor ratio $\mathscr{K}^{SP}$ compared to the competitive equilibrium, which implies a higher implicit wage $\mathscr{W}^{SP}$. 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

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

Since $\alpha < 1$, we have $\mathscr{K}^{SP} > \mathscr{K}$. 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

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

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

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

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 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

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

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

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

Define the implicit value of innovations as

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

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

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

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

□

Proposition (Restatement of Proposition 5)**.** *The social planner solution features higher technology adoptions rates for older workers and a flatter life-cycle profile of adoption thresholds, as well as a higher economic growth rate.*

Proof. The proposition follows from the results above. □

Proposition (Restatement of Proposition 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.*

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

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

From the optimal technology adoption choice if follows that

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

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. $\square$

B Data Appendix

B.1 Computer Adoption in the Workplace

All CPS data are downloaded from IPUMS (Flood et al., 2020). 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.[25]

The computer adoption measure is based on the response to the question of whether the respondent uses a computer at work. 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.[26] 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 structure. I will refer to this variable as the proficiency index.

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

[26] "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.

Table B.1: Summary Statistics for CPS Sample

Variable	**Obs.**	**Mean**	**Std. Dev.**	**Median**	**IQR**
PC Adoption	207,998	0.581	0.493	1	1
PC Proficiency	109,280	2.784	1.703	3	3
Age	207,998	41.013	9.934	40	15
Female	207,998	0.423	0.494	0	1
College Degree	207,998	0.341	0.474	0	1
Graduate Degree	207,998	0.125	0.330	0	0
White	207,998	0.846	0.361	1	0
Black	207,998	0.106	0.308	0	0
Asian	207,998	0.038	0.191	0	0

Note: This tables reports summary statistics for the CPS CIU sample. Observations are weighted by CPS CIU supplement weights. PC Adoption is an indicator variable for whether the subject uses a computer at work. PC Proficiency is an index ranging from 1 to 6 and only available when subject works with computer at work. See Appendix B for data construction details.

B.2 Data on Local Labor Markets

Census/ ACS. All data are obtained from IPUMS (Ruggles et al., 2020). I map Census geographies to commuting zones using the crosswalks provided by David Dorn on his personal data page. I winsorize all variables at the 1% and 99% level to reduce noise. I follow Acemoglu and Autor (2011) in defining full-time full-year workers as workers working at least 35 hours per week and at least 40 weeks per year.

Based on David Dorn's consistent occupational codes, 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).

Patents. All patent data is derived from the US PTO Patentsview bulk download files.[27] I map granted patents to counties based on the location of the inventor. I only consider US inventors and, in case of multiple inventors, assign each county the share of a patents according to the share of residing inventors.

I use individual citations within the first 5 years since granting to construct patent weights. Following Kogan et al. (2017), I normalizing citations by the average citation in the CPC sub-section for patents granted in the same year. This ensures that the average patent across patent classes has the same value in the aggregation. I further adjust this factor such that it aggregates to the total citations in the first five years for a given cohort. The citation adjusted value of a patent is calculated as

$$1 + \frac{\text{Citations}}{\text{Average citations in CPC-subsection}} \times \text{Average citations}$$

County level aggregates are mapped to CZs using the mapping provided on David Dorn's data page. I use a three year window average around the year for robustness. In other words, the citation weighted patents for a CZ in 1990 are its average across 1989,1990, and 1991. I further consider alternative patent weights using raw counts, unadjusted citations, citations adjusted by CPC section, as well as dropping self-citations by assignees or inventors in the baseline measure.

Fertility Instrument. The fertility instrument creates shares of working young (age 25-44) among the working-age population (age 25 to 64) by county based on fertility. To do so, I use recorded births for 1940-68, age 0 population for 1969-2016 and imputed births based on state fertility and relative fertility rates across counties for 1920-39. I detail the approach for each time-period separately in the following.

Births by county for 1920-39. There are no recorded births available for 1920-39 using either NBER Vitality data or SEER data. To impute births by county

[27] See https://www.patentsview.org/download/

I rely on Census data from IPUMS for population sizes in 1920, 1930 and 1940 by county, age 0 - 9 population at the state level and population at the state level as well as relative birth rates across counties from 1940-49.
Firstly, I impute population by county for 1921-1929 and 1931-39 by using population sizes in 1920,1930 and 1940 from the National Historical Geographic Information System (NHGIS), which is based on the Census, and assume constant geometric growth rates between census years.
Secondly, I use IPUMS USA to estimate fertility rates by year by assuming that age 0-9 population was born within each state and using geometric smoothing for state population between Census years, i.e. my estimated fertility rate for Arkansas in 1921 is the age 9 population in Arkansas in the 1930 census divided by the geometric average of 1930 and 1940 population in Arkansas based on the IPUMS USA.
Finally, I use the relative fertility rates across counties from 1940 to 1949 (see below) to allocate births within states. Births for county i in state j are then calculated as

$$Births_{it} = Pop_{it} \times Fert_{i,1940-49} \times \frac{Births_{jt}}{\sum_{i \in j} Pop_{it} \times Fert_{i,1940-49}}$$

This formula basically allocates imputed births within a state based on relative fertility rates across counties in 1940-49 and current population. Time-series variation in this measure thus arises from state-level changes in births and population movements across counties.
Births by county for 1940-68. For the 1940-69 period I rely on the "Vital Statistics Births" as provided by the NBER.[28] I do not transform the data apart from selection county-level observations and mapping county names to county FIPS codes. The latter part requires some manual remapping of counties due to changing county definitions over the historical period. The main culprits for this are Alaska and Virginia. I drop Alaska in my analysis (due to overall inconsistent data across source) and generally map independent cities

[28] Heidi Williams is credited with the compilation and, in accordance with the terms of use, I acknowledge financial support from NIA grant P30-AG012810 through the NBER.

in Virginia (the main source of inconsistencies) to the county they are formed from.

Births by county 1968 - 2016. For all other birth data I use age 0 population by county from the Surveillance, Epidemiology, and End Results program (SEER), which is provided by the NBER as well. As with the 1940-68 data, I map independent cities in Virginia to their surrounding county.

Mapping Counties to Commuting Zones. I map county-level births to 1990 commuting zones using the mapping provided by Eckert et al. (2018), which is based on the geographic shape files of the Census. The mapping changes over time and accounts for changes in the geographic definition of counties. The mapping thus allows me to approximate the births within the 1990 boundaries of commuting zones. The mapping is many to many and I aggregate using the provided weights.

Table B.2: Summary Statistics for Local Labor Markets Sample

Variable	Observations	Mean	Std. Dev.	Median	IQR
Δ R&D emp. (%)	2,166	0.428	0.969	0.360	1.263
Δ Private R&D emp. (%)	2,166	0.473	1.062	0.334	1.458
Δ Non-academic R&D emp. (%)	2,166	0.412	0.959	0.345	1.248
Δ R&D emp. in tradables (%)	2,166	0.583	1.239	0.470	1.374
Δ Citation adj. patents p.c. (CPC-sub-section)	1,955	0.711	1.227	0.542	1.520
Δ Patents p.c.	1,955	0.050	0.130	0.021	0.153
Δ Citation adj. patents p.c.	1,955	0.676	1.184	0.529	1.542
Δ Citation adj. patents p.c. (CPC-section)	1,955	0.693	1.203	0.485	1.608
Δ Citation adj. patents p.c. (CPC-sub-section)	1,955	0.711	1.227	0.542	1.520
Δ Ext. citation adj. patents p.c. (CPC-sub-section)	1,955	0.436	0.870	0.308	1.089
Δ Wage of R&D worker	2,166	11.422	8.723	11.541	9.475
Δ Private R&D Wage (%)	2,162	11.500	9.169	11.941	9.582
Δ Non-academic R&D Wage (%)	2,166	11.576	8.816	11.499	9.686
Δ Tradable R&D Wage (%)	2,152	11.673	11.449	12.366	10.836
Δ WYS	2,166	-2.713	5.979	-4.869	11.220
Δ WYS (instr.)	2,166	-5.397	4.933	-6.813	9.505
Δ WAPA	2,166	0.737	1.143	1.215	1.942
Δ WAPA (instr.)	2,166	1.175	1.020	1.411	1.938

Note: This table reports summary statistics for the local labor markets sample. Δ refers to decadal differences. Observations are weighted by 1980 population. R&D employment refers to the share of FTFY workers employed in R&D occupations. Patents per capita refers to patents per 1000 capita weighted by citations in the first 5 years after granting. WYS refers to the working young share and WAPA refers to the average working-age population age. See Appendix B for data construction details.

C Empirical Appendix

C.1 Computer Adoption by Workers

Table C.1: Regression Table for Computer Use At Work

Independent Variable	(1) **Computer Adoption (%)**	(2) **Computer Adoption (%)**	(3) **Computer Adoption (%)**	(4) **Computer Adoption (%)**	(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)
Year FEs	Yes	Yes	Yes	Yes	Yes
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. All standard errors clustered at industry level.

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

Table C.2: Regression Table for Tasks Performed With Computer

Independent Variable	(1) **Computer Proficiency**	(2) **Computer Proficiency**	(3) **Computer Proficiency**	(4) **Computer Proficiency**	(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)
Year FEs	Yes	Yes	Yes	Yes	Yes
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. All standard errors clustered at industry level.

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

Table C.3: Age vs Cohort Horserace for Computer Use At Work

Independent Variable	(1) **Computer Adoption (%)**	(2) **Computer Adoption (%)**	(3) **Computer Adoption (%)**	(4) **Computer Adoption (%)**	(5) **Computer Adoption (%)**
Age 10-14 in 1989	-2.153*	-2.626**	-1.767**	-1.835**	-1.785**
	(1.286)	(1.085)	(0.877)	(0.870)	(0.850)
Age 15-19 in 1989	0.422	-0.625	-0.493	-0.508	-0.236
	(0.805)	(0.686)	(0.534)	(0.531)	(0.558)
Age 20-24 in 1989	0.121	-0.534	-0.345	-0.350	-0.160
	(0.575)	(0.508)	(0.437)	(0.434)	(0.458)
Age 30-34 in 1989	1.272**	0.859*	0.304	0.272	0.255
	(0.585)	(0.480)	(0.378)	(0.374)	(0.351)
Age 35-39 in 1989	1.549*	0.009	-0.504	-0.581	-0.652
	(0.913)	(0.748)	(0.600)	(0.594)	(0.580)
Age 40-44 in 1989	3.205**	0.964	-1.144	-1.221	-1.212
	(1.235)	(1.023)	(0.788)	(0.778)	(0.760)
Age 45-49 in 1989	0.189	-1.003	-3.704***	-3.738***	-3.584***
	(1.539)	(1.399)	(1.061)	(1.056)	(1.020)
Age 50-54 in 1989	-2.247	-2.338	-5.770***	-5.790***	-5.750***
	(1.696)	(1.628)	(1.274)	(1.267)	(1.263)
Age 55-59 in 1989	-6.665***	-6.825***	-10.702***	-10.663***	-10.318***
	(2.123)	(2.036)	(1.608)	(1.602)	(1.557)
Age 60-64 in 1989	-10.371***	-11.138***	-14.856***	-14.751***	-13.532***
	(2.855)	(2.805)	(2.475)	(2.472)	(2.322)
Age 30-34	0.710	0.748	0.176	0.134	0.187
	(0.591)	(0.566)	(0.428)	(0.423)	(0.425)
Age 35-39	0.633	0.976	-0.209	-0.242	0.119
	(0.739)	(0.671)	(0.570)	(0.565)	(0.566)
Age 40-44	-0.294	0.032	-1.108	-1.125	-0.802
	(0.933)	(0.896)	(0.745)	(0.740)	(0.710)
Age 45-49	0.244	0.279	-0.985	-0.991	-0.803
	(1.283)	(1.167)	(0.978)	(0.976)	(0.938)
Age 50-54	-1.250	-0.947	-2.036*	-2.062*	-1.984*
	(1.683)	(1.394)	(1.163)	(1.157)	(1.147)
Age 55-59	-1.321	-0.799	-1.247	-1.230	-1.163
	(1.860)	(1.628)	(1.298)	(1.284)	(1.289)
Age 60-64	-2.749	-2.502	-1.884	-1.828	-1.890
	(2.504)	(2.196)	(1.683)	(1.677)	(1.682)
Year FEs	Yes	Yes	Yes	Yes	Yes
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 with standard deviation 49.3 and mean 58.34. Age 25 and Age 25 in 1989 are the leave out categories. Regressions

Table C.4: Age vs Cohort Horserace for Tasks Performed With Computer

Independent Variable	(1) **Computer Proficiency**	(2) **Computer Proficiency**	(3) **Computer Proficiency**	(4) **Computer Proficiency**	(5) **Computer Proficiency**
Age 10-14 in 1989	0.020	0.011	0.077	0.074	0.067
	(0.056)	(0.053)	(0.051)	(0.051)	(0.053)
Age 15-19 in 1989	0.046	0.021	0.058**	0.057*	0.059**
	(0.031)	(0.032)	(0.029)	(0.029)	(0.029)
Age 20-24 in 1989	-0.006	-0.023	-0.004	-0.004	-0.001
	(0.030)	(0.030)	(0.029)	(0.029)	(0.030)
Age 30-34 in 1989	-0.043	-0.039	-0.038*	-0.040*	-0.044*
	(0.026)	(0.025)	(0.023)	(0.023)	(0.024)
Age 35-39 in 1989	-0.095***	-0.119***	-0.110***	-0.110***	-0.117***
	(0.031)	(0.027)	(0.028)	(0.028)	(0.027)
Age 40-44 in 1989	-0.063	-0.096**	-0.111***	-0.114***	-0.121***
	(0.045)	(0.042)	(0.042)	(0.042)	(0.043)
Age 45-49 in 1989	-0.103*	-0.105*	-0.141***	-0.141***	-0.156***
	(0.057)	(0.055)	(0.051)	(0.050)	(0.050)
Age 50-54 in 1989	-0.167**	-0.144**	-0.189***	-0.194***	-0.197***
	(0.070)	(0.068)	(0.066)	(0.066)	(0.067)
Age 55-59 in 1989	-0.098	-0.085	-0.143	-0.146	-0.161*
	(0.101)	(0.098)	(0.096)	(0.095)	(0.094)
Age 60-64 in 1989	-0.188	-0.160	-0.193	-0.196	-0.191
	(0.131)	(0.129)	(0.126)	(0.125)	(0.124)
Age 30-34	0.111***	0.095***	0.085***	0.084***	0.069***
	(0.024)	(0.023)	(0.020)	(0.020)	(0.020)
Age 35-39	0.147***	0.146***	0.130***	0.131***	0.132***
	(0.033)	(0.033)	(0.031)	(0.030)	(0.031)
Age 40-44	0.133***	0.137***	0.126***	0.126***	0.132***
	(0.049)	(0.047)	(0.042)	(0.042)	(0.044)
Age 45-49	0.100*	0.104*	0.087*	0.086*	0.090*
	(0.060)	(0.058)	(0.051)	(0.051)	(0.052)
Age 50-54	0.092	0.092	0.067	0.067	0.066
	(0.077)	(0.075)	(0.066)	(0.066)	(0.066)
Age 55-59	-0.042	-0.049	-0.060	-0.057	-0.054
	(0.091)	(0.087)	(0.075)	(0.075)	(0.075)
Age 60-64	-0.169	-0.198*	-0.193**	-0.191**	-0.186*
	(0.117)	(0.109)	(0.094)	(0.094)	(0.095)
Year FEs	Yes	Yes	Yes	Yes	Yes
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 and Age 25 in 1989 are the

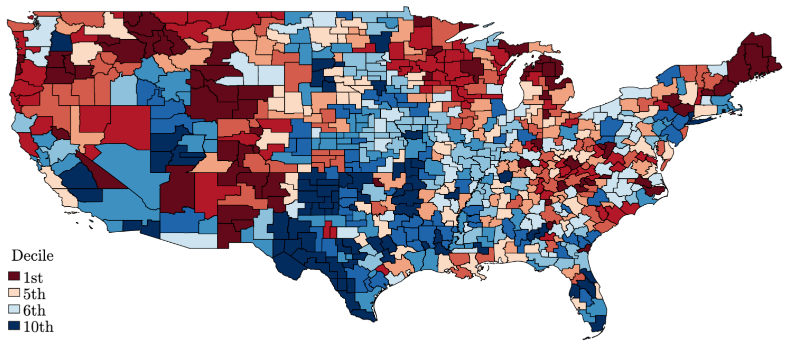
C.2 Local R&D Employment and Patenting

C.2.1 Further Discussion of the Instrument

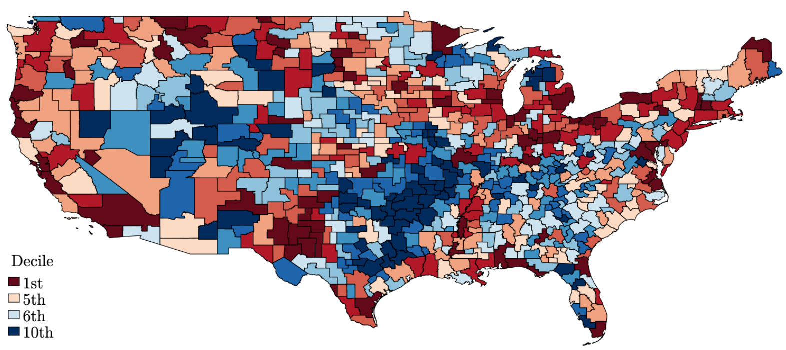
One potential source of variation are transitory economic shocks during the birth years. Schaller et al. (2020) provide evidence that birth rates respond positively to local income growth giving. This would potentially create variation in birth rates captured by the instrument. Note, however, that the variation captured would have to be transitory due to the inclusion of CZ fixed effects. Furthermore, I would only capture shock at a distant past from the actual period in consideration. Arguably, this should then be considered as quasi exogenous from the perspective of current outcomes.
Figure C.1 reports realized and instrumented changes in the WYS across US commuting zones. The figure highlights that the instrument's power is not driven by large tech-hubs such as the Bay Area or Boston. In particular, while both experienced less than average workforce aging, the instrument actually predicts the opposite.

Figure C.1: Changes in the Working Young Share Across US 1980 - 2010

A. Realized Changes



B. Instrumented Changes



Notes: Cumulative changes from 1980 to 2010 categorized into deciles. See Appendix B for data construction details.

C.2.2 Additional Results

Table C.5: Local R&D Employment and Patenting: Alternative Mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Δ R&D emp.	**Δ R&D emp.**	**Δ R&D emp.**	**Δ R&D emp.**	**Δ Patents p.c.**	**Δ Patents p.c.**	**Δ Patents p.c.**	**Δ Patents p.c.**
Δ WYS (pred.)	0.207***	0.222***	0.206***	0.203***	0.254**	0.257**	0.256**	0.258**
	(0.048)	(0.048)	(0.043)	(0.045)	(0.097)	(0.098)	(0.098)	(0.097)
Δ WA pop.		0.319***	0.196***	0.221***		0.079	0.065	0.043
		(0.102)	(0.044)	(0.049)		(0.072)	(0.075)	(0.077)
Δ College (%)			0.199***	0.198***			0.023*	0.024*
			(0.015)	(0.015)			(0.012)	(0.012)
Δ Non-white (%)				-0.008				0.007
				(0.008)				(0.008)
Δ Female (%)				-0.055				0.052
				(0.043)				(0.064)
Fixed Effects	CZ Year	CZ Year	CZ Year	CZ Year	CZ Year	CZ Year	CZ Year	CZ Year
Observations	2,166	2,166	2,166	2,166	1,924	1,924	1,924	1,924

Note: This table reports the coefficient estimates for specification (24) for R&D employment and patenting. Columns (1)-(4) report results for R&D employment shares, while (5)-(8) report the results for citation-weighted patents. Δ refers to decadal differences. WA refers to the working age population in logs. LF refers to the labor force participation rate among age 25 to 64 subjects. College degree, non-white, and female refer to the respective shares among age 25-64 population. Observations are weighted by 1980 working-age population and standard errors clustered at the state level.

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

Table C.6: Local Patenting: Alternative Measures

	(1)	(2)	(3)	(4)	(5)
	Δ Patents p.c.				
Δ WYS	0.019**	0.205*	0.202**	0.254**	0.179***
	(0.009)	(0.108)	(0.098)	(0.100)	(0.059)
Adjustments	None	Citations	Citations by CPC Section	Citations by CPC Sub-Section	External citations by CPC Sub-Section
Fixed Effects	CZ Year	CZ Year	CZ Year	CZ Year	CZ Year
F statistic (1st)	1,924	1,924	1,924	1,924	1,924

Note: This table reports the IV coefficient estimates for specification (23) for alternative patent based measures. Column (1) reports results for patents counts, while column (2) adds citations. Column (3) and (4) normalize citations counts by respective group means within a cohort, while column (5) counts only drops citations by the same assignee or inventor. See Appendix B for data construction details. The reported F-statistic is the Sanderson-Windmeijer multivariate F-test of excluded instruments. Observations are weighted by 1980 working-age population and standard errors clustered at the state level.

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

Table C.7: Local R&D Employment: Alternative Measures

	(1)	(2)	(3)	(4)
Employment		**Δ R&D emp.**		
Δ WYS	0.207***	0.242***	0.208***	0.031
	(0.049)	(0.050)	(0.049)	(0.069)
Annual wages		**Δ R&D wages**		
Δ WYS	1.707***	1.305**	1.571***	0.189
	(0.518)	(0.606)	(0.523)	(0.920)
Sample	Baseline	Private sector	Non-academic	Tradables
Fixed	CZ	CZ	CZ	CZ
Effects	Year	Year	Year	Year
F statistic (1st)	49.0	49.3	49.0	49.6
Observations	2,166	2,162	2,166	2,152

Note: This table reports the IV coefficient estimates for specification (23) for R&D employments among Full-Time Full-Year employees ans well as their log annual wages. The top panel reports results for employment, while the bottom panel reports results for wages. Column (1) reports the baseline results for referece. Columns (2)-(4) restrict the set of employees considered to private sector, non-academic, and tradable employment respectively. See Appendix B for data construction details. The reported F-statistic is the Sanderson-Windmeijer multivariate F-test of excluded instruments. Observations are weighted by 1980 working-age population and standard errors clustered at the state level.

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

Table C.8: Local R&D Employment and Patenting: Weighting Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ R&D emp.	**Δ R&D emp.**	**Δ R&D emp.**	**Δ Patents p.c.**	**Δ Patents p.c.**	**Δ Patents p.c.**
Δ WYS	0.207***	0.183***	0.220***	0.254**	0.244***	0.252***
	(0.049)	(0.059)	(0.046)	(0.100)	(0.085)	(0.092)
Weight	1980 Working-age population	Unweighted	Working-age population	1980 Working-age population	Unweighted	Working-age population
Fixed	CZ	CZ	CZ	CZ	CZ	CZ
Effects	Year	Year	Year	Year	Year	Year
F statistic (1st)	49.0	17.4	42.2	49.2	15.8	42.0
Observations	2,166	2,166	2,166	1,924	1,924	1,924

Note: This table reports the second stage IV coefficient estimates for specification (23) using alternative observations weights. Columns (1) and (4) report the baseline results, while (2) and (5) use no observation weights, and (3) and (6) use beginning of period working age population weights. See Appendix B for data construction details. The reported F-statistic is the Sanderson-Windmeijer multivariate F-test of excluded instruments. Observations are weighted by 1980 working-age population and standard errors clustered at the state level.

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

Table C.9: Local R&D Employment and Patenting: Sources of Variation

	(1)	(2)	(3)	(4)	(5)	(6)
	Δ R&D emp.	**Δ R&D emp.**	**Δ R&D emp.**	**Δ Patents p.c.**	**Δ Patents p.c.**	**Δ Patents p.c.**
Δ WYS	0.210***	0.321***	0.846**	0.268***	0.386**	1.012*
	(0.051)	(0.094)	(0.413)	(0.097)	(0.146)	(0.507)
Fixed	CZ	CZ	CZ	CZ	CZ	CZ
Effects	Year	Year × Region	Year × State	Year	Year × Region	Year × State
F statistic (1st)	52.8	23.5	4.8	53.0	23.7	5.0
Observations	2,157	2,157	2,157	1,915	1,915	1,915

Note: This table reports the second stage IV coefficient estimates for specification (23) adding geography-year interactions. Columns (1) and (4) report the baseline results, while (2) and (5), and (3) and (6) add region-year and state-year fixed effects. See Appendix B for data construction details. The reported F-statistic is the Sanderson-Windmeijer multivariate F-test of excluded instruments. Observations are weighted by 1980 working-age population and standard errors clustered at the state level.

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