The Price of Experience*

Hyeok Jeong†
KDI School of Public Policy and Management

Yong Kim‡
Yonsei University

Iourii Manovskii§
University of Pennsylvania

July 15, 2014

Abstract

We identify a key role of factor supply, driven by demographic changes, in shaping several empirical regularities that are a focus of active research in macro and labor economics. In particular, demographic changes alone can account for the large movements of the return to experience over the last four decades, for the differential dynamics of the age premium across education groups emphasized by Katz and Murphy (1992), for the differential dynamics of the college premium across age groups emphasized by Card and Lemieux (2001), and for the changes in cross-sectional and cohort-based life-cycle profiles emphasized by Kambourov and Manovskii (2005).

*We would like to thank Daron Acemoglu, Rui Castro, Aureo De Paula, Jeremy Greenwood, Elena Krasnokutskaya, Bob Lucas, Richard Rogerson, Ludo Visschers and seminar participants at the Atlanta Fed, U of Alberta, UC Berkeley - Haas School of Business, U of Bonn, UC San Diego, UC Santa Barbara, Cleveland Fed, Korea University, NYU, NYU - Stern School of Business, Seoul National University, U of Pennsylvania, the Philadelphia Fed, the Fed Board of Governors, USC - Marshall School of Business, U of Toronto, 2008 Korean Econometric Society Summer Camp Macroeconomics Meeting, the 2007 NBER Summer Institute, New York Fed /Philadelphia Fed Conference on Quantitative Macroeconomics, 2008 Midwest Macro meetings, 2008 Minnesota Workshop on Macroeconomic Theory, 2008 SED annual meeting, and the Search & Matching Club meeting at the University of Pennsylvania for their comments. This research has been supported by the National Science Foundation Grants No. SES-0617876, SES-0922406 and the Bank of Korea.

†KDI School of Public Policy and Management, 87 Hoegiro Dongdaemun-gu, Seoul 130-868, Korea. E-mail: hyeokj@gmail.com.

‡School of Economics, Yonsei University, Daewoo Hall 509, 262 Seongsan-no, Seodaemun-gu, Seoul 120-749, Korea. E-mail: yongkim75@gmail.com.

§Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA, 19104-6297 USA. E-mail: manovski@econ.upenn.edu.

1 Introduction

We assess the role of supply, driven by the demographic changes of the workforce, in accounting for some of the prominent and widely studied US labor market trends. In particular, we find that the large movement of the rate of return to experience over the 1968-2007 period can be nearly perfectly explained by demographic changes alone, with no role attributable to demand shifts. Moreover, these demographic changes account for the differential dynamics of the age premium across schooling groups emphasized by Katz and Murphy (1992), for the differential movements of the college premium across age groups emphasized by Card and Lemieux (2001), and for the changes in cohort-based life-cycle earnings profiles documented in Kambourov and Manovskii (2005). Thus, our analysis attributes a key role to demographic change in shaping several empirical regularities that are a focus of active research in macro and labor economics.

Our modeling approach is based on the seminal analysis of the role of supply and demand factors in driving the experience differentials in Katz and Murphy (1992). These authors assumed that workers supply two distinct productive inputs to the aggregate production. In particular, they assume that young workers (within 5 years from completing schooling) supply exclusively one of these inputs, which we will refer to as "pure labor," or just "labor." Old workers (25-35 years after schooling completion) are assumed to supply exclusively the other input, which we will refer to as the "pure experience," or simply "experience." Other workers are assumed to supply a bundle of these two inputs. The amount of labor and experience supplied by the workers from a particular demographic cell is determined by projecting their wages on wages of workers from the two age groups exclusively supplying "pure" inputs. The correlation between the relative wages of the age groups exclusively supplying "pure" inputs and the relative aggregate supplies of experience and labor then identifies the aggregate elasticity of substitution between these two factors.

We retain the key modeling elements but change the measurement strategy. In particular, we do not assume that old workers completely stop supplying one input and supply exclusively the other. Instead, we allow older workers to provide both inputs and determine the quantities they supply using the variation in actual number of years worked among individuals of a given age.[1] We implement this strategy by measuring the effective amount of labor and experience supplied by each worker through decomposing individual wages using a version of the classic Mincerian wage equation. This provides a natural way to control for exogenous changes to the returns of other productive attributes (schooling, gender, race etc.) when measuring the aggregate relative quantity and price of experience.[2] The individual wage equations are consistent with the aggregate production function and aggregate to determine the total supplies of effective labor and experience.

The specification of the wage equation is designed to obtain a good fit to individual wages. It places no restriction on whether the dynamics of the relative price of experience is driven by the relative supply or the relative demand for experience. Nevertheless, our empirical approach reveals a correlation between the aggregate relative price of experience and aggregate relative supply of experience of -0.95 over our forty year sample period. We find no role for the demand shifts in accounting for the dynamics of the relative price of experience.

Our use of the panel data drawn from the Panel Study of Income Dynamics (PSID) allows us to identify how the effective supplies of labor and experience vary with age. This allows us to study the role of the changing relative price of experience in shaping the dynamics of the age premium (the relative wage of young and old workers). For any demographic subgroup, the wage response to a change in the aggregate price of experience is determined by the share of wages sourced from experience. Thus, as the remuneration for experience accounts for a larger share of wages of older workers, the age premium responds positively to the price of experience. However, the strength of this response can vary across demographic groups depending on how

[1] Appendix A1 presents a simple example illustrating the differences between the two approaches.

[2] This presents a challenge to the demographic cell-based correction for composition that represents the current state-of-the-art in this literature. For example, Katz and Murphy (1992) construct wages for a potential experience group by adding up the wages of all demographic subgroups belonging to it using the fixed weights given by the average employment shares of the subgroups. This procedure controls for changes in composition of other attributes, e.g., education and gender, however, it cannot filter out the effects of changing returns to other attributes. For instance, changes in the college premium will differentially affect the wages of pure young and pure old workers, and this will bias the estimates of the relative price of experience. Similarly, the labor and experience supplies are aggregated using fixed efficiency weights, whereas changes in the college premium etc. should be reflected in the measurement of effective supplies.

much their effective stock of experience relative to labor changes with age. We show that this insight quantitatively reconciles an economy-wide movement in the price of experience with the differential movements in the age premium across schooling groups.

This is relevant for assessing the role of supply and demand factors because the age premium rose much more in the 1970s and 80s for non-college than for college educated workers. The approach in Katz and Murphy (1992) cannot rationalize the differential movement in age premiums across demographic groups in response to the change in the aggregate relative supply of experience, because their empirical strategy does not differentiate between experience and age (the old and the young are assumed to exclusively supply experience and labor, respectively). This led them and the subsequent literature to interpret this evidence as representing a shift in demand against young high school educated men. In contrast, we find no role for such demand effects.

There is an important corollary to this finding. In an influential article, Card and Lemieux (2001) show that the college premium rose sharply after 1973 for young workers while it remained more or less constant for old workers. Thus, the relative college premium fell for old workers. Since the ratio of the age premiums between college and non-college workers is mathematically equivalent to the ratio of the college premiums between old and young workers, this finding is a dual of the finding in Katz and Murphy (1992) that the relative age premium fell for college workers. Thus, the fact that we match the evolution of the age premium across schooling groups implies that we also account for the differential movement of the college premium across age groups.[3] Thus, we identify the changing supply of experience, induced largely by the labor force entry of the baby boom cohorts and women, as the common source of changes in the aggregate relative price of experience, in the age premium across demographic groups, and the cohort effects in the college premium which have previously been studied separately from each other.

While much of the literature focuses on cross-sectional changes in life-cycle earnings profiles, a life-cycle profile of a cohort of workers, say, entering the labor market at the same time is perhaps more relevant for the decision making of members of that cohort. In a stationary environment

[3] As in Card and Lemieux (2001), we do not attempt to explain the changes in the economy-wide college premium which are treated as exogenous, and are the subject of debate.

cross-sectional and cohort-based profiles coincide. When the relative price of experience changes, however, cross-sections are no longer useful guides to life-cycle earnings. In fact, Kambourov and Manovskii (2005) have noted that a substantial steepening of the cross-sectional profiles in the 1970s and 1980s was accompanied by flattening of life-cycle earnings profiles for successive cohorts of male workers entering the labor market in the 1970s and 1980s. This has potentially large macroeconomic consequences. For example, Elsby and Shapiro (2012) use the flattening of cohort-based life-cycle profiles to account for changing labor market participation rates, while Song and Yang (2012) use it to understand changes in the savings rates. However, the cross-sectional and cohort-based profiles are intimately related. Consequently, we show that the flip side of our model's ability to account for the dynamics of cross-sectional earnings is its ability to account for the changes in cohort-based profiles. Thus, we find the demographic change that governs the evolution of the relative price of experience is also responsible for the changes in cohort-based earnings profiles.

The remainder of the paper is organized as follows. In Section 2, we provide an overview of our measurement approach. In Section 3, we provide descriptive analysis that, without imposing the model structure, illustrates the very strong relationship between the relative price and the relative supply of experience. In Section 4, we structurally estimate the model and evaluate the role of supply and demand factors in accounting for the evolution of the relative price of experience over time. In Section 5, we use the estimated model to study the role of the dynamics of the supply of experience in driving (1) age premium across education groups, (2) college premium across age cohorts, (3) the rates of return to years of prior work in the aggregate and across demographic groups, (4) changing cohort-based life-cycle earnings profiles, and (5) the decline in the entry wages in the 1970s and 1980s and their subsequent increase as well as the measured total factor productivity. Section 6 concludes.

2 From Individual Earnings to Aggregate Quantities: Overview of the Measurement Approach

2.1 Individual Earnings

Consider an individual i who at date t supplies $\hat{l}_{it}$ units of effective labor input and $\hat{e}_{it}$ units of effective experience input to the labor market. The individual works for h_{it} hours and has idiosyncratic productivity z_{it}. If market prices of labor and experience inputs are given by R_{Lt} and R_{Et}, individual earnings can be written as

$$y_{it} = \left[R_{Lt}\hat{l}_{it} + R_{Et}\hat{e}_{it}\right] z_{it}h_{it} \equiv R_{Lt}\left[\hat{l}_{it} + \Pi_{Et}\hat{e}_{it}\right] z_{it}h_{it}, \tag{1}$$

where $\Pi_{Et} \equiv \frac{R_{Et}}{R_{Lt}}$ denotes the relative price of experience to labor. This implies the log-wage equation

$$\ln w_{it} = \ln R_{Lt} + \ln\left[\hat{l}_{it} + \Pi_{Et}\hat{e}_{it}\right] + \ln z_{it}, \tag{2}$$

that, as we show below, can be extended to be suitable for empirical work and estimated to recover the individual stocks of effective labor and experience and the aggregate time series of the relative price of experience.

2.2 Aggregate Technology

We consider an aggregate production function that maps the aggregate stock of labor L_t and the aggregate stock of experience E_t into aggregate labor earnings Y_t,

$$Y_t = A_t G\left(L_t, E_t\right), \tag{3}$$

where G is a constant-returns-to-scale function and A_t represents the aggregate productivity of the composite input of experience and labor.[4] We assume G is continuous and differentiable in its arguments, and the Euler theorem implies $Y_t = A_t\left(G_{L_t}L_t + G_{E_t}E_t\right)$, where $G_{L_t} = \frac{\partial G}{\partial L_t}$

[4] More precisely, let $\tilde{A}$ denote the Hicks-neutral productivity affecting the constant-returns-to-scale aggregate production function for output $\tilde{Y} = \tilde{A}F(K,G)$, that takes the capital stock K and the the composite input of labor and experience $G(L,E)$ as inputs. Then the aggregate labor earnings are given by $Y = \tilde{Y} - \tilde{A}F_K\left(\frac{K}{G}\right)K = \tilde{A}F_G\left(\frac{K}{G}\right)G \equiv AG$, Thus, the tfp term affecting aggregate labor earnings is given by $A \equiv \tilde{A}F_G\left(\frac{K}{G}\right)$.

and $G_{E_t} = \frac{\partial G}{\partial E_t}$ will be referred to as marginal products of labor and experience (net of the productivity A_t), respectively.

Competitive firms can bundle workers to maintain the desired experience to labor ratio as in Heckman and Scheinkman (1987). This implies that prices of the two services provided by workers are competitively determined:[5]

$$R_{L_t} = A_t G_{L_t}, \tag{4}$$

$$R_{E_t} = A_t G_{E_t}. \tag{5}$$

Then, the relative price of experience $\Pi_{Et} = \frac{G_{E_t}}{G_{L_t}}$, is falling in the ratio of aggregate experience to labor $\frac{E_t}{L_t}$, as long as $G_{E_t L_t} > 0$: that is, as long as experience and labor are complements.

2.3 Consistent Aggregation

Summing the individual earnings equation in (1) over individuals i at a given date t, we have

$$\begin{aligned} \sum_i y_{it} &= R_{L_t} \sum_i \hat{l}_{it} z_{it} h_{it} + R_{E_t} \sum_i \hat{e}_{it} z_{it} h_{it} \\ &= A_t G_{L_t} L_t + A_t G_{E_t} E_t \\ &= Y_t \end{aligned}$$

where the aggregate inputs L_t, E_t are measured as

$$L_t = \sum_i \hat{l}_{it} z_{it} h_{it}, \tag{6}$$

$$E_t = \sum_i \hat{e}_{it} z_{it} h_{it}, \tag{7}$$

and prices R_{L_t}, R_{E_t} are determined by Equations (4) and (5), respectively.

Thus, the individual earnings in equation (1) *consistently aggregate* to the aggregate earnings as is implied by the aggregate production function in (3). This consistent aggregation holds for any homogeneous of degree one function G as long as the aggregate inputs L_t and E_t are consistently measured as in equations (6) and (7).

[5] The aggregate production function approach (with the competitive pricing of the bundled inputs) has been recently used in the literature to study earnings dynamics, e.g., Heckman, Lochner, and Taber (1998) and Guvenen and Kuruscu (2009, 2010).

2.4 Two Approaches to Estimating the Aggregate Production Function Parameters

To conduct quantitative analysis one must choose a specific functional form for G. We follow Katz and Murphy (1992) and restrict our attention to the commonly used class of constant elasticity of substitution (CES) production technologies. Specifically,

$$Y_t = A_t \left(L_t^{\mu} + \delta E_t^{\mu} \right)^{\frac{1}{\mu}}, \tag{8}$$

where the elasticity of substitution between L_t and E_t is measured by $\frac{1}{1-\mu}$ (where $\mu \leq 1$), and the parameter $\delta > 0$ adjusts the relative scale between L_t and E_t. The degree of substitutability between experience and labor is governed by the value μ. If $\mu = 1$, experience and labor are perfect substitutes. In this case, the demographic change affecting the ratio of labor to experience does not affect their relative price. However, if $\mu < 1$ so that labor and experience are not perfect substitutes, changes in the demographic composition of the workforce will affect the relative price of experience. This aggregate production function implies that the relative price of experience is given by

$$\Pi_{Et} = \delta \left(\frac{E_t}{L_t} \right)^{\mu - 1}. \tag{9}$$

To estimate the parameters δ and μ we will follow two approaches.

1. The estimation of the log wage equation in (2) delivers the time-series for the relative price of experience Π_{Et} and the measure of effective labor and experience at the individual level, which can be added up as in equations (6) and (7) to obtain the aggregate inputs L_t and E_t. Taking logs on both sides of equation (9), one immediately verifies that a simple regression of the time-series of log relative price of experience on a constant and the time-series of the log of relative supplies allows one to estimate the parameters δ and μ. This procedure does not impose the restrictions implied by the functional form of the production function on the individual earnings equations. Thus, the measurement of the aggregate prices and quantities is independent of the particular form of the aggregate production function. Consequently, this procedure is in no way hardwired to obtain an estimate of μ that is statistically different from 1. If changes in relative prices are independent from

changes in relative supplies, the regression will reveal this. We pursue this approach in Section 3.

2. An alternative approach that imposes full model structure on individual earnings equations and is therefore more efficient is to substitute the explicit expression for the relative price of experience given by equation (9) into the individual log wage equation (2) and estimate its parameters directly from the micro data (of course, maintaining the consistent aggregation by imposing equations (6) and (7)). This procedure is computationally considerably more demanding but it also places no ex-ante restrictions making it likely to find evidence of complementarity between the two inputs. We pursue this approach in Section 4.

Finally, to assess the role of changes in aggregate demand driving the relative price of experience in either of the two approaches, δ can be assumed to be a particular function of time as in Katz and Murphy (1992), and the parameters of this function can be estimated together with μ.

3 Descriptive Analysis

In this section, we (1) extend the specification of the individual earnings equations to make them suitable for the empirical work, (2) estimate the time-series of the aggregate relative price of experience and of the aggregate supplies of labor and experience using the PSID data, and (3) obtain preliminary estimates of the aggregate production function parameters. The estimation in this section does not impose the specification of the aggregate production function on the individual earnings equations. Thus, the resulting estimate of the aggregate relative price of experience is not restricted by the theory, and we will be able to assess the performance of the theory by its ability to match this time series.

3.1 Measuring the Aggregate Prices and Quantities of Labor and Experience using Micro Data

Our objective in modeling individual earnings is to develop a specification such that the mapping from the individual characteristics to wages picks up the first order features of wage differentials as documented by the vast existing literature. The basis of our empirical specification is the traditional Mincer equation that is a cornerstone of the empirical work in labor economics.[6] This specification has been remarkably successful at empirically describing individual earnings with only a few shortcomings highlighted in recent work (e.g., Heckman, Lochner, and Todd (2006), Lemieux (2006)). A slight but economically interesting "fine-tuning" of the equation allows us to overcome these shortcomings.[7]

Recall that our basic log-wage equation is given by

$$\ln w_{it} = \ln R_{Lt} + \ln \left[\hat{l}_{it} + \Pi_{Et} \hat{e}_{it} \right] + \ln z_{it}.$$

We follow the traditional Mincerian specification and model the individual productivity variable $\ln z_{it}$ as determined by the vector χ_{it} of the observable characteristics affecting log earnings in an additive manner:

$$\ln z_{it} = \alpha_t \chi_{it}, \tag{10}$$

where the vector χ_{it} includes the observable characteristics, such as years of schooling, sex, race and geographic region, and α_t is the vector of associated coefficients.

The only difference between our specification and the Mincerian one is that the standard Mincerian specification assumes that there is only one productive input supplied by workers. In contrast, following Katz and Murphy (1992), we seek to assess the possibility that workers supply two distinct factors to aggregate production and whose aggregate relative supplies affect their relative prices. If individual earnings consist of a sum of payments to two distinct factors of production, the logarithm of earnings is not equal to the sum of logarithms of constituent factors. This explains the presence of the $\ln \left[\hat{l}_{it} + \Pi_{Et} \hat{e}_{it} \right]$ term. Our specification, however nests the traditional one.

In modeling the life-cycle curvature we follow the approach pioneered by Katz and Murphy (1992). They divided workers into demographic cells based on age, gender, etc. Ignoring, for

[6] Wasmer (2001a) also considers a Mincerian wage equation approach to link the return to years worked to the aggregate relative supply of experience.

[7] There is considerable debate in the literature regarding the interpretation of the Mincerian coefficient on education. While Mincer (1958, 1974) provided two sets of assumptions that allow for a particular structural interpretation of this coefficient, these assumptions are often questioned in the literature (e.g., Heckman, Lochner, and Todd (2006)). We do not insist on such an interpretation.

the ease of exposition, the partition other than age, consider N groups of workers $j_1, j_2, ..., j_N$. Katz and Murphy assumed that young workers in group j_1 exclusively supply the labor input while old workers in group j_N exclusively supply the experience input. But how much labor and experience do all other age groups supply? They answer this question by assuming the existence of time invariant schedules $\lambda_L(j)$, $\lambda_E(j)$ that measure the quantity of the two inputs supplied by workers of age j. They measured these supplies by regressing wages of age j workers on wages of workers in groups j_1 and j_N (without intercept term). The coefficients from these regressions are taken as the measure of the two inputs entering the aggregate production function supplied by workers of age j.

Our framework relies on the same schedules λ_L and λ_E,[8] although our measurement approach is different because we do not impose the assumption that old workers exclusively supply the experience input. Instead, these schedules are identified from the variation in the number of years actually worked or "years worked," e_{it}, among workers of a given age as discussed in Appendix A1.[9]

Specifically, the life-cycle profile depends on age through time invariant age efficiency schedules for labor and experience $\lambda_L(j_{it}, s_{it}, x_{it})$ and $\lambda_E(j_{it}, s_{it}, x_{it})$, where j denotes age measured in years, $s_{it} \in \{HS, C\}$ denotes either a "high school" education group HS with years of schooling less than or equal to 12, or a "college" education group C with years of schooling beyond 12, and $x_{it} \in \{M, F\}$, denotes gender with M being male and F being female.[10] This implies that, aside from individual productivity z_{it}, the effective supply of labor of individual i at date t, $\hat{l}_{it}$,

[8] In particular, we maintain the assumption that these schedules are exogenously given. Our objective in this paper is to explore the effects of changing only the measurement approach while retaining the structure of the key papers in the literature, such as Katz and Murphy (1992), Krusell, Ohanian, Ríos-Rull, and Violante (2000), Card and Lemieux (2001), etc. Thus, we leave an explicit optimizing behavioral modeling of these profiles for important future work.

[9] In Appendix A3 we show that our model is non-parametrically identified if individuals of the same age differed by only one year of prior work. We also document there that the amount of variation available in the data vastly exceed this sufficient condition, even among male workers (as was also documented by Light and Ureta (1995), among others).

[10] A number of articles, e.g., Heckman, Lochner, and Todd (2006), have found that the standard Mincerian specification that implies that earnings profiles (based on years since completing schooling) are parallel across demographic groups is at odds with the data. Allowing the relationship between wages and age and experience to differ across schooling and gender groups addresses this shortcoming.

depends on age, gender, and education, and is given by $\lambda_L(j_{it}, s_{it}, x_{it})$. Similarly, the effective supply of experience, $\hat{e}_{it}$, depends on years worked, age, gender, and education, and is given by $\lambda_E(j_{it}, s_{it}, x_{it})g(e_{it})$.[11] Note that we allow the technology for accumulating experience input to exhibit curvature. In other words, the mapping from years of actual prior work e_{it} to accumulated effective units of experience is given by some function $g(e_{it})$, not necessarily linear. We show in Appendix A6 that all these sources of curvature are essential for fitting the wage distribution. It is convenient to define the relative age efficiency schedule of experience as the ratio of the age efficiency schedule of experience to that of labor $\lambda_{E/L}(j_{it}, s_{it}, x_{it}) \equiv \frac{\lambda_E(j_{it}, s_{it}, x_{it})}{\lambda_L(j_{it}, s_{it}, x_{it})}$.

With these definitions, the log wage equation can be written as

$$\ln w_{it} = \ln R_{Lt} + \ln \lambda_L \left(j_{it}, s_{it}, x_{it}\right) + \ln \left[1 + \Pi_{Et} \lambda_{E/L} \left(j_{it}, s_{it}, x_{it}\right) g\left(e_{it}\right)\right] + \ln z_{it}. \quad (11)$$

For the empirical implementation we adopt the following parsimonious but flexible specifications which approximate the age efficiency schedules $\lambda_L(j_{it}, s_{it}, x_{it})$ and $\lambda_E(j_{it}, s_{it}, x_{it})$ by an exponential function of a second-degree polynomial with coefficients that are allowed to vary with gender and education

$$\begin{aligned} \lambda_L(j_{it}, s_{it}, x_{it}) &= \exp(\lambda_{L,0}\left(s_{it}, x_{it}\right) + \lambda_{L,1}\left(s_{it}, x_{it}\right) j_{it} + \lambda_{L,2}\left(s_{it}, x_{it}\right) j_{it}^2), \quad (12) \\ \lambda_E(j_{it}, s_{it}, x_{it}) &= \exp(\lambda_{E,0}\left(s_{it}, x_{it}\right) + \lambda_{E,1}\left(s_{it}, x_{it}\right) j_{it} + \lambda_{E,2}\left(s_{it}, x_{it}\right) j_{it}^2). \quad (13) \end{aligned}$$

This implies that the relative age efficiency schedule of experience is given by

$$\lambda_{E/L}(j_{it}, s_{it}, x_{it}) = \exp(\lambda_{E/L,0}\left(s_{it}, x_{it}\right) + \lambda_{E/L,1}\left(s_{it}, x_{it}\right) j_{it} + \lambda_{E/L,2}\left(s_{it}, x_{it}\right) j_{it}^2), \quad (14)$$

where $\lambda_{E/L,k}\left(s, x\right) = \lambda_{E,k}\left(s, x\right) - \lambda_{L,k}\left(s, x\right)$ for $k \in \{0, 1, 2\}$. Without loss of generality, we normalize $\lambda_{L,0}\left(s, x\right) = \lambda_{E,0}\left(s, x\right) = 0$ for the low education ($s = HS$) group.

For the technology mapping years worked to experience input, we assume a flexible quartic specification

$$g\left(e_{it}\right) = e_{it} + \theta_1 e_{it}^2 + \theta_2 e_{it}^3 + \theta_3 e_{it}^4. \quad (15)$$

[11] To facilitate the transparency of some derivations below we do not allow for the differences in parameters of this mapping across schooling and gender groups. We have verified that allowing for such heterogeneity has no substantive impact on any of the results in the paper. Introducing the direct dependence of g on age is also inconsequential.

Substituting expressions (10), (12), (14), and (15) into equation (11), and substituting $\ln R_{L_t}$ by a time dummy D_t we obtain the log wage equation to be estimated

$$\begin{aligned}\ln w_{it} \;=\;& D_t + \left(\lambda_{L,0}\left(s_{it},x_{it}\right)+\lambda_{L,1}\left(s_{it},x_{it}\right)j_{it}+\lambda_{L,2}\left(s_{it},x_{it}\right)j_{it}^2\right) \qquad (16)\\ &+\ln\left[1+\Pi_{E_t}\exp\left(\begin{array}{c}\lambda_{E/L,0}\left(s_{it},x_{it}\right)\\ +\lambda_{E/L,1}\left(s_{it},x_{it}\right)j_{it}+\lambda_{E/L,2}\left(s_{it},x_{it}\right)j_{it}^2\end{array}\right)\left(\begin{array}{c}e_{it}+\theta_1 e_{it}^2\\ +\theta_2 e_{it}^3+\theta_3 e_{it}^4\end{array}\right)\right]\\ &+\alpha_t\chi_{it}+\epsilon_{it},\end{aligned}$$

where ϵ_{it} represents a classical measurement error.

Estimating equation (16), we obtain the estimates of parameters of the efficiency schedules $\widehat{\lambda}_L(j,s,x)$, $\widehat{\lambda}_E(j,s,x) = \widehat{\lambda}_{E/L}(j,s,x)\widehat{\lambda}_L(j,s,x)$ and $\widehat{g}\left(e\right)$, as well as the estimates of the time-varying coefficients ($\widehat{\alpha}_t$, $\widehat{D}_t$, and $\widehat{\Pi}_{E_t}$). These estimates allow us to construct the estimated aggregate labor and experience inputs $\widehat{L}_t$ and $\widehat{E}_t$ at each date t as

$$\widehat{L}_t = \sum_i \widehat{\lambda}_L(j_{it},s_{it},x_{it})\widehat{z}_{it}h_{it}, \qquad (17)$$

$$\widehat{E}_t = \sum_i \widehat{\lambda}_E(j_{it},s_{it},x_{it})\widehat{g}\left(e_{it}\right)\widehat{z}_{it}h_{it}. \qquad (18)$$

To investigate the role of supply in determining the evolution of prices, we seek to document a relationship between the estimated relative price $\widehat{\Pi}_{Et}$ and the estimated relative supply of experience $\frac{\widehat{E}_t}{\widehat{L}_t}$ constructed using these equations. We should emphasize that we did not impose any of the model structure on these earnings equations (except for perfect competition and constant returns to scale in aggregate production). Thus, there is no hardwired relationship between $\widehat{\Pi}_{Et}$ and $\frac{\widehat{E}_t}{\widehat{L}_t}$. To the extent that we find a relationship between them, it will be suggestive of the nature of the aggregate production function.

3.2 Results

We obtain estimates of the parameters applying a nonlinear least-squares method to the log-wage equation (16). The estimates of these coefficients and their standard errors are reported in Appendix Tables A-1 and A-2. In what follows, we first discuss the implied aggregate quantities of interest, followed by the discussion of the individual wage determination.

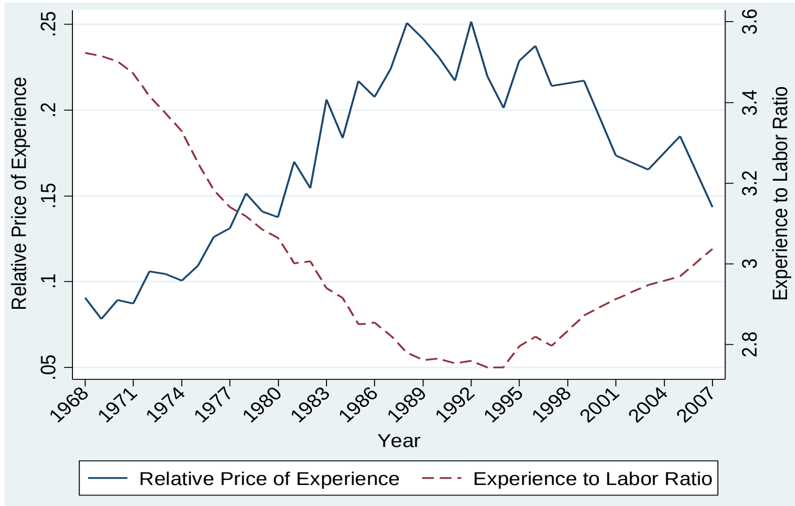


Figure 1: The Estimated Relative Price of Experience $\widehat{\Pi}_{E_t}$ and the Experience to Labor Ratio $\frac{\widehat{E}_t}{\widehat{L}_t}$.

3.2.1 The Relative Supply of Experience and its Relative Price

The estimated series of the relative price of experience $\widehat{\Pi}_{E_t}$ and the implied aggregate experience-labor ratio $\frac{\widehat{E}_t}{\widehat{L}_t}$ are plotted in Figure 1. We refer to these estimates as "unrestricted" as they are independent of the aggregate production function we are ultimately interested in estimating. There is a substantial movement of the relative price of experience, which increases with an average growth rate of 5.1% per year between 1968-1988 and falls thereafter back to its level in 1979 by 2007. The estimated experience-labor ratio over the same period displays a clear negative co-movement with the relative price of experience. The correlation coefficient between the log relative price of experience and log of the experience-labor ratio is remarkably high at -0.95. Thus, the unrestricted data imply a very strong co-movement between the relative supply of experience and its relative price.[12] This suggests that the aggregate technology features complementarity between aggregate experience and labor. Next, we obtain preliminary estimates of this technology.

[12]The analysis based on Katz and Murphy's sample and methodology delivers a correlation between the relative wage and relative supplies of 0.6 to 0.8.

Table 1: Estimates of Technology Parameters based on Unrestricted Estimates of the Relative Price of Experience $\widehat{\Pi}_{E_t}$ and the Experience to Labor Ratio $\frac{\widehat{E}_t}{\widehat{L}_t}$.

parameter	Benchmark (1)	With Demand Shifts (2)
μ	-3.21	-3.35
	(0.07)	(0.24)
δ	16.82	20.16
	(1.28)	(5.91)
δ_1	—	-0.03
		(0.04)
R^2	0.92	0.99

Note - Entries for μ, δ and the goodness-of-fit represent the results of regressing the unrestricted relative price of experience on a constant and the relative supply of experience. For sample restrictions and variable construction procedures, see Appendix A2.

3.2.2 Preliminary Estimation of the Aggregate Production Function Parameters

Taking logs on both sides of equation (9) which determines the relative price of experience given the CES aggregator of labor and experience, we obtain:

$$\ln \Pi_{Et} = \ln \delta + (\mu - 1) \ln \left(\frac{E_t}{L_t} \right). \tag{19}$$

Treating this equation as a regression of the unrestricted relative price of experience on a constant and the unrestricted relative supply of experience to labor, one can obtain estimates of the parameters δ and μ. These estimates are summarized in Column (1) of Table 1. The results of this experiment suggest that the simple aggregate production function that features complementarity between the aggregate supplies of labor and experience can rationalize the movements in the relative price of experience remarkably well through the changing relative supply of experience.

To assess whether the changing demand for experience played a role in determining its relative price, we re-estimate Equation (19) by allowing the intercept to be a linear function of time, i.e., $\delta_t = \delta + \delta_1 t$. This is the standard approach to assessing the role of demand shifts in the literature. We find that the estimate of δ_1 is statistically insignificant. Moreover, allowing for the demand trend does not affect the estimate of the complementarity parameter μ. This is in contrast to the findings based on the empirical approach of Katz and Murphy (1992). Repeating the same experiment using their data and empirical methodology one finds that the relative price of experience is entirely driven by the change in demand with the estimate $\widehat{\mu}$ statistically indistinguishable from one.[13]

While standard in the literature, allowing for only a linear demand trend might appear restrictive, especially given that the relative price of experience follows an inverted-U pattern. On the other hand, a sufficiently flexible specification of the exogenous demand shifts can always mechanically fit the dynamics of the relative wages. Thus, without a deeper theory on the nature of the demand shifts and their measurement we have no guidance on how they should be modeled.[14] In contrast, a simple model based on the observed dynamics of supply is clearly successful in accounting for the dynamics of the relative price of experience.

3.2.3 Sources of Curvature in Life-Cycle Wage Profiles

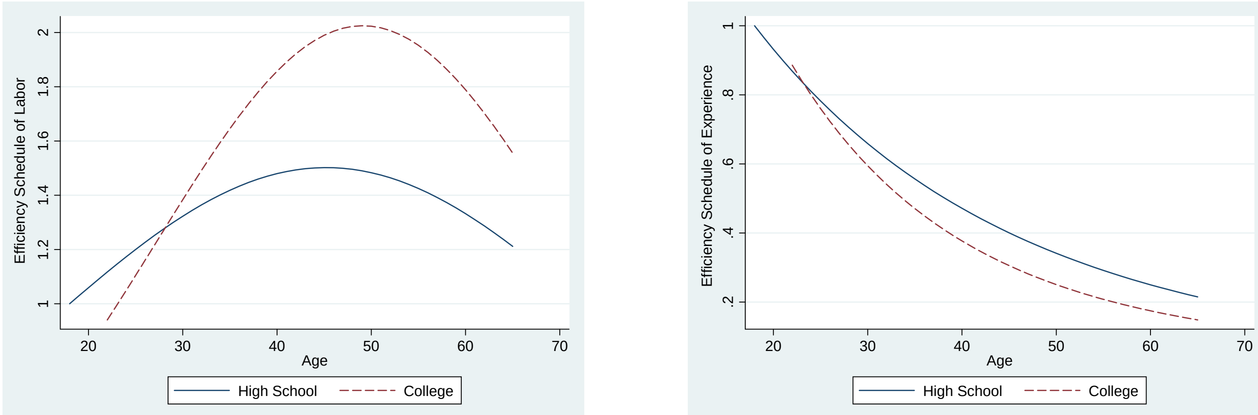
While the key features of the age efficiency schedules are the same for both genders, we focus the discussion on the efficiency schedules of male workers as these are relevant for the results reported below. The estimated age efficiency schedules of experience and labor for male workers from the two schooling groups are presented in Figure 2. The corresponding estimated schedules for female workers can be found in Appendix Figure A-3. The coefficient estimates and their standard errors reported in Appendix Table A-1 imply that the efficiency schedules are estimated very precisely.

Figure 2(a) illustrates that the male age efficiency schedules of labor, $\lambda_L(j_{it}, s_{it}, M)$, are hump-shaped, peaking at age 50 for the college group and age 45 for the high school group. The increase in the effective units of labor in the early part of the life-cycle is considerably larger for the college-educated workers than their high-school educated counterparts.

In contrast, the male estimated age efficiency schedules for experience $\lambda_E(j_{it}, s_{it}, M)$ are

[13]Weinberg (2005) also emphasizes the role of demand effects.

[14]We nevertheless experimented with allowing for a quadratic trend in demand $\delta_t = \delta + \delta_1 t + \delta_2 t^2$ and found that this specification also yields highly statistically insignificant estimates of δ_1 and δ_2, and has no impact on the estimate of the complementarity parameter μ or the fit of the regression.



(a) Age Efficiency Schedule of Labor.

(b) Age Efficiency Schedule of Experience.

Figure 2: Estimated Age-Efficiency Schedules for Male Workers by Education.

monotonically decreasing and are below unity over the entire age range as shown in Figure 2(b). This suggests a substantial benefit from accumulating experience early on in life, and that this benefit is larger for college-educated workers, since their stock of experience depreciates at a faster rate. These observations imply that the male relative age efficiency schedule $\lambda_{E/L}(j_{it}, s_{it}, M)$ is falling in age and is uniformly lower for the college schooling group.

Meanwhile, the mapping from years worked to the stock of experience input via $g\left(e_{it}\right)$ is close to being linear (although mildly concave), implying that the hump shape in life-cycle wage profile is not driven by the decreasing returns to accumulating experience. Note that the effective supply of experience is the product of the age efficiency schedule for experience and the stock of experience, implying that for a typical worker the effective supply of experience increases over most of the life-cycle.

In Appendix Figure A-2 we plot the age-conditional means and standard deviations of wages and experience (both years of prior work and the amount of experience input, $g(e)$, accumulated by an individual) for each gender/education combination. We find that the age-conditional mean experience grows almost linearly with age, although there are differences in the slope across the subgroups, while the age-conditional mean wage grows in a concave manner over age. This helps explain our finding that the rate at which total experience translates into higher wages declines rapidly with age.[15]

[15] We thank an anonymous referee for suggesting this figure.

By re-estimating the model on various time subsamples or even on particular cross-sections of the data we found that the shapes and even the magnitudes of the estimated age efficiency schedules are very robust. This suggests that they are primarily identified by the heterogeneity of the number of years actually worked at a given age and not by the time variation in the data. In Appendix A5 we present evidence supporting the restriction that age efficiency schedules are independent of time.

3.2.4 Estimation on Male Sample

So far we have measured the relative prices and quantities on the same samples. While it is clear that factor supplies by female workers must be included in aggregate quantities of labor and experience, one may by concerned with the possibility that strong changes in female participation patterns over time may induce selection effects that may bias the estimates of the dynamics of the relative price of experience.

To assess the robustness of our findings to allowing for this possibility we estimate the individual earnings equations on the sample of male workers only. The resulting estimates generate the time series of the relative price of experience $\widehat{\Pi}_{E_t}^{M}$ which is plotted in Figure 3. As it is estimated on a substantially smaller sample, this estimated relative price of experience series is more volatile, but its dynamics are similar to those estimated on the full sample (which are plotted in Figure 1).

The preliminary estimates of technology are little affected by this change in the measurement procedure with the estimated elasticity of substitution between labor and experience declining slightly form 0.24 in the benchmark to 0.15 in this experiment. The dynamics of the relative price of experience continues to be well explained by the aggregate supply of experience. Allowing for a time trend in the parameter δ to capture the potential role of demand shifts continues to yield highly insignificant estimates of the associated parameters. We conclude that the possible selection biases associated with estimating the returns to experience on a panel of female workers have at best a minor impact on our findings.

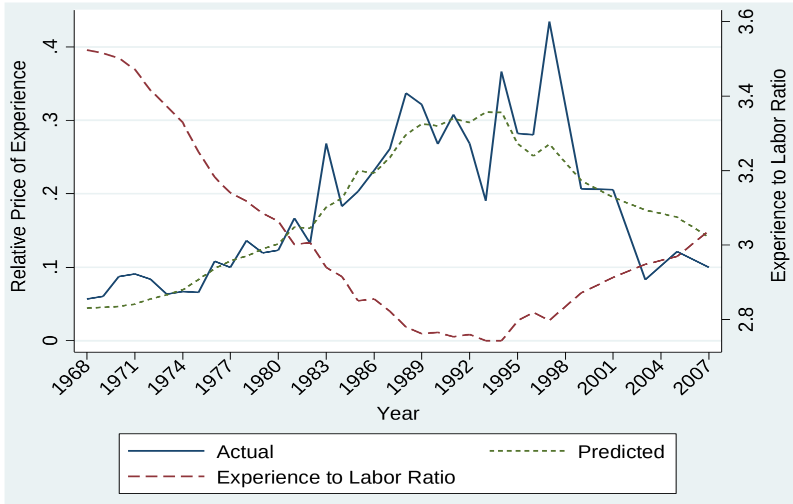


Figure 3: The Unrestricted Relative Price of Experience $\widehat{\Pi}_{E_t}$ Estimated on the Male Sample and the Predicted Male Relative Price of Experience from the Benchmark Structural Model.

4 Structural Estimation Imposing Full Model Structure

In this section, we conduct a structural estimation of the model by imposing all the properties of the aggregate production function on the individual earnings equations. This is a more efficient approach to estimating the aggregate production function parameters. However, imposing the full model structure may result in changes to the estimates of all the parameters and in the associated fit of the earnings equations. Indeed, the structural model imposes strong additional restrictions and has considerably less flexibility in fitting the wage data as compared to the unrestricted specification studied above. The extent to which a structural model can rationalize the relationship between variables of interest without sacrificing the fit to the data provides a measure of the quality of the model.

4.1 Measuring the Aggregate Prices and Quantities of Labor and Experience using Individual Data

To structurally estimate the model, we modify the specification of the earnings equation estimated in Section 3 by substituting the unrestricted relative price of experience with the expression implied by the aggregate production function.

$$
\begin{aligned}
\ln w_{it} \quad = \quad & D_t + \left(\lambda_{L,0}\left(s_{it}, x_{it}\right) + \lambda_{L,1}\left(s_{it}, x_{it}\right) j_{it} + \lambda_{L,2}\left(s_{it}, x_{it}\right) j_{it}^2\right) \\
& + \ln\left[1 + \delta\left(\frac{E_t}{L_t}\right)^{\mu-1} \exp\begin{pmatrix} \lambda_{E/L,0}\left(s_{it}, x_{it}\right) \\ +\lambda_{E/L,1}\left(s_{it}, x_{it}\right) j_{it} \\ +\lambda_{E/L,2}\left(s_{it}, x_{it}\right) j_{it}^2 \end{pmatrix} \begin{pmatrix} e_{it} + \theta_1 e_{it}^2 \\ +\theta_2 e_{it}^3 + \theta_3 e_{it}^4 \end{pmatrix}\right] \\
& + \alpha_t \chi_{it} + \epsilon_{it},
\end{aligned} \tag{20}
$$

where consistent aggregation requires that the aggregate inputs L_t and E_t are measured as in equations (17) and (18) and hence depend on the parameters of efficiency schedules $\lambda_L(j, s, x)$, $\lambda_E(j, s, x)$, $g(e)$ and the coefficient vector α_t. That is, for a consistent estimation of the entire model, the aggregate inputs L_t and E_t need to be expressed in terms of these life-cycle efficiency and productivity parameters and estimated at the same time. This creates a complicated nonlinearity inside the arguments of the already nonlinear log wage equation.

To estimate the model, we use the following iterative guess-and-verify strategy. Guess the parameters for $\lambda_L(j, s, x)$, $\lambda_E(j, s, x)$, $g(e)$ and α_t and compute the implied aggregate experience to labor ratio. Using this guessed ratio, estimate all parameters of the log wage equation using a non-linear least squares method. Then, verify if the estimates for the parameters for $\lambda_L(j, s, x)$, $\lambda_E(j, s, x)$, $g(e)$ and α_t from this estimation coincide with the initial guess. If not, recalculate the experience-labor ratio using the obtained estimates of those parameters, and iterate this procedure until the guessed estimates and the subsequent estimates coincide.

Identification of the aggregate production function parameters is discussed in Appendix A7.

4.2 Results

The estimates of the parameters of Equation (20) and their standard errors are reported in Appendix Tables A-4 and A-5. All estimates are very similar to those obtained without imposing the aggregate production function on individual wage equations. In fact, the fit of the model remains unchanged. As discussed above, this indicates the appropriateness of our modeling choices.

The structural estimates of the technology parameters are reported in Column (1) of Table 2. These estimates continue to imply fairly strong complementarity between labor and experience

Table 2: Structural Estimates of Technology Parameters.

parameter	Benchmark (1)	With Demand Shifts (2)
μ	-2.95	-3.02
	(0.23)	(0.28)
δ	13.26	14.57
	(3.45)	(4.87)
δ_1	—	-0.02
		(0.03)
R^2	0.92	0.92
$RMSE$	0.62	0.62

Note - Entries for μ, δ and the goodness-of-fit represent the results of structural estimation. For sample restrictions and variable construction procedures, see Appendix A2.

in aggregate production. To assess the role of the changes in the aggregate relative supply of experience in driving its relative price, we plot in Figure 4 the predicted relative price of experience based on the structural estimates and the unrestricted relative price of experience obtained in Section 3. Despite the parsimonious specification (two parameters μ and δ, and a single state variable $\frac{E}{L}$), the model tracks the actual time-path of the relative price of experience very closely. The correlation coefficient between the unrestricted relative price of experience $\widehat{\Pi}_{E_t}$ and the predicted one based on the structural estimates is 0.97.

The tight prediction of the relative price of experience generated by the experience-labor ratio in Figure 4 leaves very little room for the demand based explanations. To investigate the role of demand shifts more formally, we allow for the share parameter δ in the production function (8) to vary over time as $\delta_t = \delta + \delta_1 t$. The corresponding estimates for the technology parameters are reported in Column (2) of Table 2. The fit of the model is the same as in the benchmark specification, and the predicted relative price of experience is indistinguishable across the specifications. The estimate of δ_1 is economically and statistically insignificant. The estimate of the complementarity parameter μ is virtually the same as in the benchmark specification.

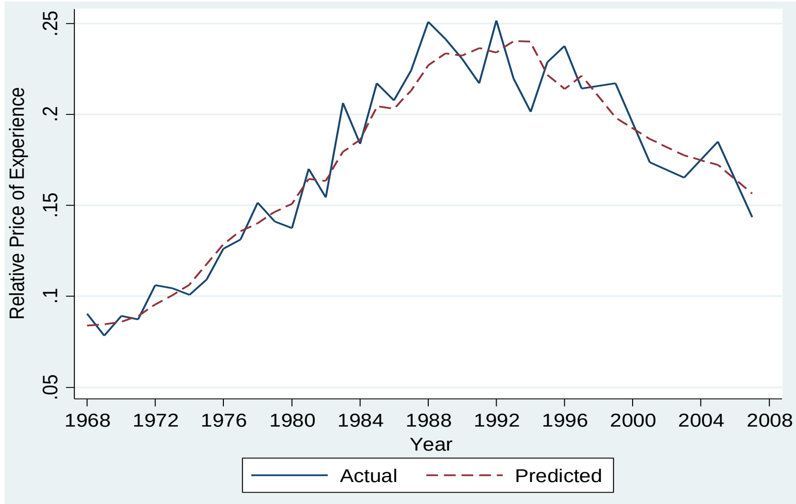


Figure 4: The Unrestricted Relative Price of Experience $\widehat{\Pi}_{E_t}$ and the Predicted Relative Price of Experience from the Benchmark Structural Model.

5 Using the Model to Understand Additional Dimensions of Wage Dynamics

In this section, we use the estimated model to study the role of the dynamics of the supply of experience in driving (1) age premium across education groups, (2) college premium across age cohorts, (3) the rates of return to years of prior work in the aggregate and across demographic groups, (4) changing cohort-based life-cycle earnings profiles, and (5) the measured total factor productivity.

5.1 Age Premium across Schooling Groups

Katz and Murphy (1992) were among the first to find that the ratio of wages of old and young male workers, or age premium, has increased more among high school educated workers than among college educated workers in the 1980s. Our PSID data exhibit the same patterns. The lines labeled "Actual" in Figure 5 represent the ratio of wages of "old" male workers aged 41-60 to wages of "young" male workers, aged 18-40 among the college educated and the high school educated workers, respectively.[16] As discussed in the Introduction, the standard measurement

[16] We focus on two broad age categories to define the young and old to encompass the entire sample, and to minimize the small sample issues arising from using the PSID. We note however, that all the results are robust

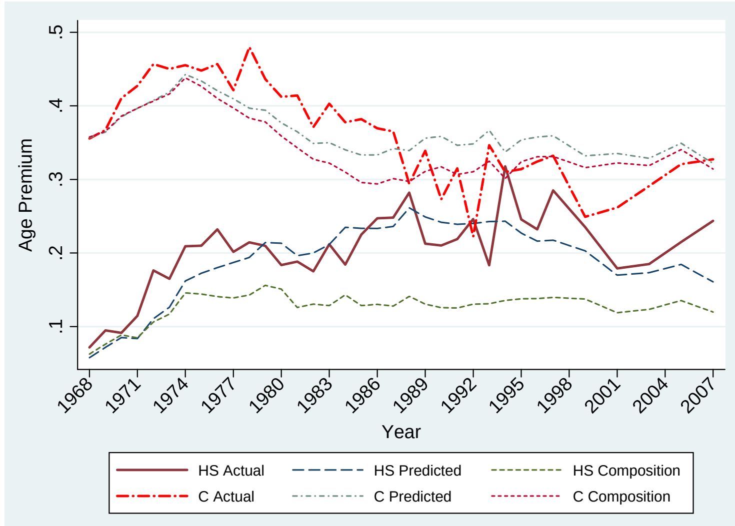


Figure 5: Age Premium by Schooling Group.

approach pioneered by Katz and Murphy (1992), assumes that the pure old exclusively supply experience and the pure young exclusively supply labor regardless of schooling group. This assumption implies that, in contrast to the data, the wage ratio w^s_{old}/w^s_{young} has an elasticity with respect to relative price of experience of one regardless of schooling group $s \in \{HS, C\}$. This led the literature to emphasize the role of demand shifts against young high school educated males.

In contrast, our model of individual earnings distinguishes age and experience, and allows the age premium to exhibit a different trend over time from the relative price of experience. Specifically, the age premium, r^s_t, among workers with schooling level s is given at date t by[17]

$$
\begin{aligned}
r^s_t \equiv \ln\left[\frac{w^s_t\,(old)}{w^s_t\,(young)}\right] &= \ln\left[\frac{\lambda_L\,(old, s)}{\lambda_L\,(young, s)}\right] \\
&+ \ln\left[\frac{1+\delta\left(\frac{E_t}{L_t}\right)^{\mu-1}\lambda_{E/L}\,(old, s)\, g\,\left(e^s_t\,(old)\right)}{1+\delta\left(\frac{E_t}{L_t}\right)^{\mu-1}\lambda_{E/L}\,(young, s)\,\; g\left(e^s_t\,(young)\right)}\right] \\
&+ \alpha_t\left[\chi^s_t\,(old) - \chi^s_t\,(young)\right].
\end{aligned}
\tag{21}
$$

Thus, the level of the age premium is determined by the differences in the stocks of effective labor and experience supplied by the young and old workers, and by the difference in the vector

to finer partitions of workers into age groups.

[17]For notational clarity we suppress the dependence of all equations in this and the next section on gender.

of their demographic characteristics χ. Over time, the age premium changes due to the evolution of the relative price of experience $\Pi_{Et} \equiv \delta \left(\frac{E_t}{L_t}\right)^{\mu-1}$, as well as the change in the demographic composition of young and old groups and changes in the "returns" to their characteristics, α_t.

Equation (21) implies that, among workers with schooling level s, the elasticity of the age premium with respect to the aggregate relative price of experience is given by

$$\epsilon_{r_t^s,\Pi_{E_t}} = \frac{\epsilon_{w_t^s,\Pi_{E_t}}(old) - \epsilon_{w_t^s,\Pi_{E_t}}(young)}{r_t^s}, \tag{22}$$

where

$$\epsilon_{w_t^s,\Pi_{E_t}}(j) = \frac{\delta \left(\frac{E_t}{L_t}\right)^{\mu-1} \lambda_{E/L}(j,s)\, g\,\left(e_t^s(j)\right)}{1 + \delta \left(\frac{E_t}{L_t}\right)^{\mu-1} \lambda_{E/L}(j,s)\, g\,\left(e_t^s(j)\right)} \tag{23}$$

is the elasticity of wages of age j workers with respect to the aggregate relative price of experience.

These equations are insightful. First note that the elasticity of wages with respect to the relative price of experience, $\epsilon_{w_t^s,\Pi_{E_t}}$, is equal to the share of wages that represents the remuneration to experience. Quite naturally, wages of workers with higher relative effective supply of experience, $\lambda_{E/L}(j,s)\, g\,\left(e_t^s(j)\right)$, respond more to changes in the price of experience. As our estimates imply that old workers generally have higher relative effective supply of experience, their wages rise more in response to an increase in the relative price of experience than wages of young workers whose earnings mostly represent the remuneration for labor. This implies that the age premium tends to co-move with the relative price of experience.

However, the strength of the response of age premium to changes in the relative price of experience across, say, schooling groups depends on how fast the share of wages sourced from experience rises with age among workers belonging to those groups. Our estimates imply that this share rises faster with age among high school educated than among college educated workers. Consequently, the age premium is more responsive to changes in the aggregate relative price of experience among high school educated workers then among their more educated counterparts.

Quantitatively we find that our model does a very good job at matching the evolution of the age premium across education groups. The lines labeled "Predicted" in Figure 5 plot the age premium for the two schooling groups implied by the structurally estimated parameters, including the technology parameters $\hat{\mu}$, $\hat{\delta}$, and the changing relative supply of experience. The lines labeled "Counterfactual" eliminate the effect of the evolution of the relative supply of experience by setting $\mu = 1$. Thus, these lines isolate the impact on the age premium of changes in composition of workers within age and schooling groups and changing "returns" to their characteristics (the $\alpha_t \left[\chi_t^s \left(old\right) - \chi_t^s \left(young\right)\right]$ term in equation (21)).

We observe that the changing relative supply of experience is the primary determinant of the dynamics of age premium among high school educated workers through its effect on the relative price of experience. The age premium is highly responsive to changes in the relative price of experience among these workers. Changes in composition play an important role in the early 1970s but have virtually no impact on the subsequent evolution of age premium in this group of workers. In contrast, as the share of wages sourced from experience is similar in our samples of young and old college educated workers, changes in the relative price of experience have little impact on age premium in this schooling group. Instead, it is the changing composition of workers that is responsible for much of the change in the age premium.

Thus, we conclude that our measurement approach that separates the effects of age and experience on wages can quantitatively rationalize the differential impact of the aggregate relative price of experience on age premium across schooling groups. It interprets this difference not as evidence of demand shifts against particular groups of workers, but through the different importance of the experience input in wages of various schooling groups. This result depends, of course, on the estimated shape of the age efficiency schedules for labor and experience. We would like to emphasize that the estimates of these efficiency schedules were not targeted to match the asymmetric movements of age premium across schooling groups. They are *time invariant* parameters, and were estimated to fit the overall wage dispersion over the entire sample period rather than the age premium dynamics itself.[18] Given this, we view the ability to match these patterns in the data as strong evidence in support of the model.

[18]Furthermore, as we mentioned above, the estimates of the age efficiency schedules are very robust to estimating them on different time subsamples of our data, or even in various cross-sections. This suggests that the dynamics of age premium across schooling groups plays a minor role in identifying the parameters of the efficiency schedules. See also Appendix A5.

5.1.1 Ratio of Age Premiums

Mathematically, the ratio of the age premiums between college and high school educated workers is equivalent to the ratio of the college premiums between old and young groups:

$$\frac{w_{old}^{C}/w_{young}^{C}}{w_{old}^{HS}/w_{young}^{HS}} \equiv \frac{w_{old}^{C}/w_{old}^{HS}}{w_{young}^{C}/w_{young}^{HS}}.$$

Thus, the fact that the model matches the evolution of the age premiums across schooling groups implies that it can also quantitatively account for the differential movement of the college premiums across age groups. In the following section, we use this insight to show that the change in the relative supply of experience that drives the movement of the relative price of experience over time, can also account for the differential changes in the returns to schooling among young and old workers.

5.2 College Premium across Age Groups

In an influential paper, Card and Lemieux (2001) have highlighted the fact that changes in the male college premium have been very different across age groups: the college premium rose sharply after the early 1970s among young workers while it remained more or less constant among old workers. The lines labeled "Actual" in Figure 6 show that the same pattern holds in the PSID data.

In our model, the college premium, r_t^j, among workers in age group j is given at date t by

$$\begin{aligned} r_t^j \equiv \ln\left[\frac{w_t^C(j)}{w_t^{HS}(j)}\right] &= \ln\left[\frac{\lambda_L(j,C)}{\lambda_L(j,HS)}\right] \\ &+ \ln\left[\frac{1+\delta\left(\frac{E_t}{L_t}\right)^{\mu-1}\lambda_{E/L}(j,C)\ g\left(e_t^C(j)\right)}{1+\delta\left(\frac{E_t}{L_t}\right)^{\mu-1}\lambda_{E/L}(j,HS)\,g\ \left(e_t^{HS}(j)\right)}\right] \\ &+ \alpha_{ys,t}\left[ys_t^C(j)-ys_t^{HS}(j)\right]+\alpha_{-ys,t}\left[\chi_{-ys,t}^{C}(j)-\chi_{-ys,t}^{HS}(j)\right]. \end{aligned} \tag{24}$$

In the bottom row of Equation (24) we split the vector of productive characteristics into the difference in years of schooling of college and high school workers in a given age group $\left[ys_t^C(j)-ys_t^{HS}(j)\right]$ multiplied by the aggregate "returns" to schooling $\alpha_{ys,t}$ and the vector of differences in other productive characteristics $\left[\chi_{-ys,t}^{C}(j)-\chi_{-ys,t}^{HS}(j)\right]$ with the associated vector of coefficients $\alpha_{-ys,t}$. This emphasizes the fact that the increase in the aggregate "returns' to

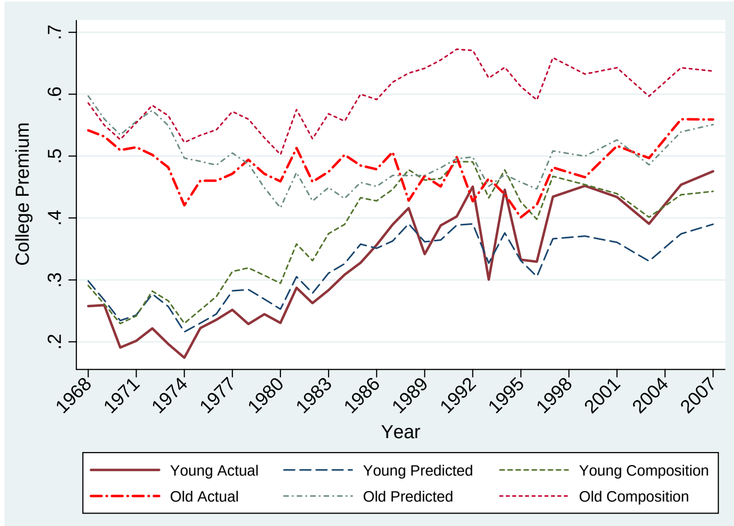


Figure 6: College Premium by Age Group.

schooling, $\alpha_{ys,t}$, symmetrically increases the college premium in all age groups. However, this effect is modulated by changes in the relative price of experience $\Pi_{Et} \equiv \delta \left(\frac{E_t}{L_t}\right)^{\mu-1}$. In particular, the elasticity of the college premium with respect to the aggregate relative price of experience among age j workers is given by

$$\epsilon_{r_t^j,\Pi_{E_t}} = \frac{\epsilon_{w_t^j,\Pi_{E_t}}(C) - \epsilon_{w_t^j,\Pi_{E_t}}(HS)}{r_t^j}, \tag{25}$$

where, as in Equation (23),

$$\epsilon_{w_t^j,\Pi_{E_t}}(s) = \frac{\delta \left(\frac{E_t}{L_t}\right)^{\mu-1} \lambda_{E/L}(j,s)\, g\left(e_t^s(j)\right)}{1 + \delta \left(\frac{E_t}{L_t}\right)^{\mu-1} \lambda_{E/L}(j,s)\, g\left(e_t^s(j)\right)} \tag{26}$$

is the elasticity of wages of age j workers in schooling group s with respect to the aggregate relative price of experience.

As we discussed in Section 5.1, high school educated workers derive a higher share of wages as payments to their supply of the experience input at almost all ages. Moreover, the share of wages sourced from experience increases faster over the life-cycle among high school educated compared to college educated workers. These observations imply that (1) an increase in the relative price of experience lowers the college premium in every age group, and (2) this effect is stronger among older workers. In other words, the college premium among old workers falls relative to young workers when the relative price of experience rises.

These results explain the success of our model in capturing the dynamics of the college premium across age groups. The lines labeled "Predicted" in Figure 6 plot the college premium for the two age groups implied by the structurally estimated parameters. The fit to the raw data series in solid lines is quite close. Comparing to the "Counterfactual" predictions that eliminate the effect of the aggregate relative supply of experience on the relative price of experience by setting $\mu = 1$, we conclude that it was the rise in the relative price of experience that almost fully counteracted the effect of the rise in the "returns" to college among old workers, yielding an essentially constant college premium among the old. As there is only a small difference in the share of wages due to experience among college and high school educated young workers, the effect of the rise in the relative price of experience played only a minor role in shaping their college premium, and was not sufficiently strong to counteract the effect of the rising "returns" to education. These results demonstrate that the changing relative supply of experience can simultaneously account for the dynamics of the relative price of experience and for the differential movements of the college premium across age groups. While the literature has treated these developments as being unrelated, our analysis establishes a very tight link between them.[19]

5.3 Rate of Return to Years Worked

While our focus in the preceding sections was on understanding the dynamics of the aggregate relative price of experience, the primary focus of the related applied literature is on understanding the evolution of the rate of return to a year of prior work, Ω, which is the coefficient on individual's number of years worked in the traditional log wage equation. In this section we provide a mapping between the two concepts. Define the rate of return Ω as the marginal wage increment to the addition of one more year of prior work:

$$\Omega_{it} \equiv \frac{d \ln w_{it}}{de_{it}} = \frac{\delta \left(\frac{E_t}{L_t}\right)^{\mu-1} \lambda_{E/L}(j_{it}, s_{it}, x_{it}) g'(e_{it})}{1 + \delta \left(\frac{E_t}{L_t}\right)^{\mu-1} \lambda_{E/L}(j_{it}, s_{it}, x_{it}) g(e_{it})}. \tag{27}$$

The individual rate of return Ω_{it} is rising in the aggregate relative price of experience $\Pi_{Et} \equiv \delta \left(\frac{E_t}{L_t}\right)^{\mu-1}$, and falling in the individual level of e_{it} (given that $g(e_{it})$ is mildly concave). The return is also increasing in the relative efficiency schedule $\lambda_{E/L}(j_{it}, s_{it}, x_{it})$ which, according to

[19] Carneiro and Lee (2011) discuss age premia in the context of the Card and Lemieux (2001) analysis.

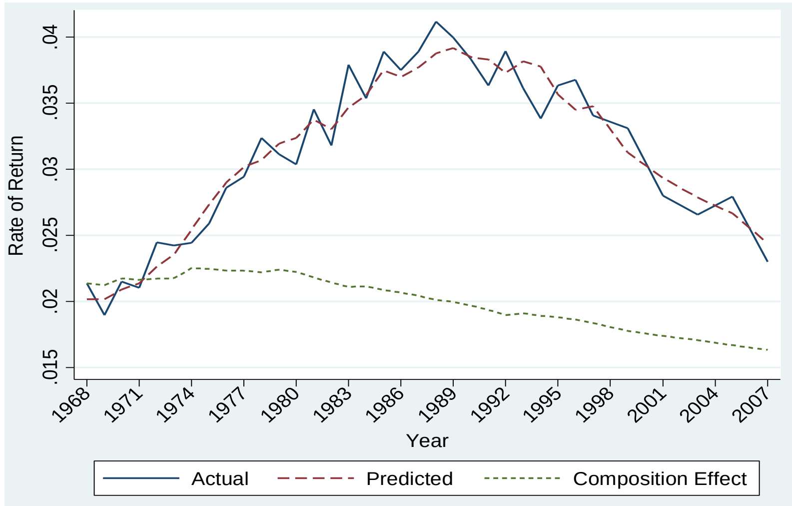


Figure 7: Actual and Predicted Rates of Return to Years Worked for the Representative Worker and Counterfactual with Constant Relative Price of Experience.

our estimates, is higher for younger than for older workers and for high school than for college educated workers. Thus, the rate of return to years of prior work declines with age and education.

5.3.1 Dynamics of Aggregate Ω

To summarize the evolution of the rate of return to years worked in a single time-series, we consider a "representative worker" whose rate of return Ω is given by $\frac{\Pi_{E_t}\Lambda_t}{1+\Pi_{E_t}\Xi_t}$ where

$$\Lambda_t \equiv \frac{\sum_i \lambda_E(j_{it}, s_{it}, x_{it}) z_{it} h_{it} g'(e_{it})}{\sum_i \lambda_L(j_{it}, s_{it}, x_{it}) z_{it} h_{it}}, \qquad \Xi_t \equiv \frac{\sum_i \lambda_E(j_{it}, s_{it}, x_{it}) z_{it} h_{it} g(e_{it})}{\sum_i \lambda_L(j_{it}, s_{it}, x_{it}) z_{it} h_{it}}. \tag{28}$$

Hence, this worker supplies the aggregate effective labor and aggregate effective experience.[20]

The solid line, labeled "Actual," in Figure 7, represents the rate of return for the representative worker constructed using unrestricted estimates of $\widehat{\Pi}_{E_t}$, $\widehat{\lambda}_L(j,s,x)$, $\widehat{\lambda}_E(j,s,x)$, $\widehat{g}(e)$, and $\widehat{\alpha}_t$ obtained in Section 3. The rate of return, measured directly in the data without imposing the model structure, is sizable and changes substantially over time from 2.1% in 1968 to 4.1% in 1988, and then back down to 2.3% in 2007. The dotted line, labeled "Predicted," is the predicted path of the rate of return implied by the structurally estimated parameters in Section

[20]By construction, the effective units of experience to labor for this representative worker coincide with the aggregate experience to labor ratio $\Xi_t = \frac{E_t}{L_t}$. We weight the age efficiency schedules of experience and labor by $z_{it}h_{it}$ to obtain the effective units of experience and labor at the aggregate level.

4, including the technology parameters $\hat{\mu}$, $\hat{\delta}$, and the changing relative supply of experience. The fit is clearly very good.[21]

Equation (27) implies that there are two potential sources of change in Ω_{it} for workers with given numer of year of prior work e_{it}: first, changes in the aggregate relative price of experience $\delta\left(\frac{E_t}{L_t}\right)^{\mu-1}$, and second, changes in the demographic composition of workers affecting the measured Λ_t and Ξ_t. To isolate the contribution of the latter effect, we set $\mu = 1$ to eliminate the complementarity effect, and normalize δ to match the level of the relative price of experience in 1968. With these restrictions, we generate a counterfactual series of the rate of return driven entirely by the composition effect (labeled "Composition Effect" in Figure 7). Clearly, the composition effect alone does not account for much of the observed changes in the aggregate rate of return to years worked. Instead, it is the changing relative supply of aggregate experience coupled with the complementarity of labor and experience in aggregate production that drives the changes in the aggregate rate of return to years of prior work.

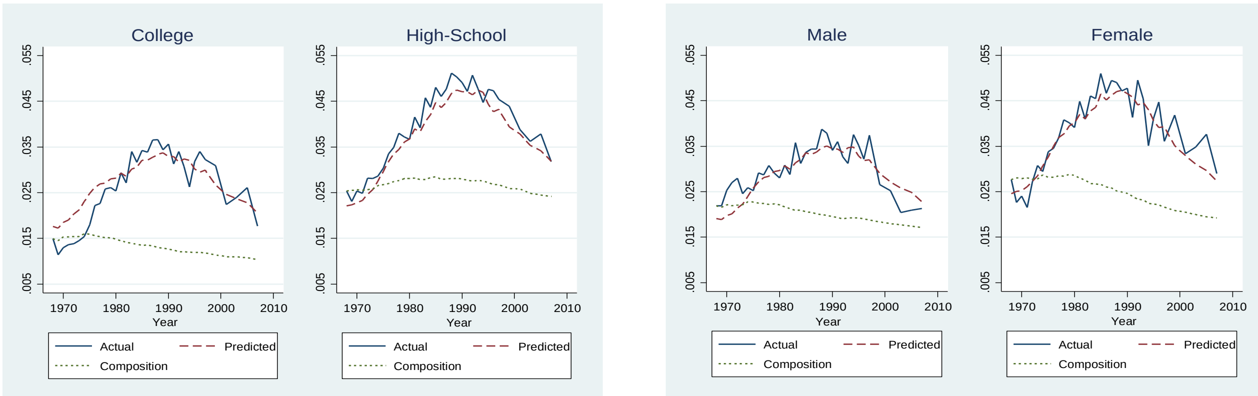
5.3.2 Rates of Return to Years Worked for Demographic Subgroups

In this section we study whether our model is consistent with the evidence on the differential response of the returns to years of prior work of various demographic groups.

To establish an appropriate benchmark we re-estimate the unrestricted specification of Section 3 of the paper but now allow the parameter Π_{E_t} to be specific to gender and education groups and evolve *independently* across these groups (i.e., we replace Π_{E_t} in equation (16) with $\Pi_{E_t}(s_{it}, x_{it})$). Thus, we effectively estimate all the parameters and the implied path of the rate of return to years worked, $\Omega(s_{it}, x_{it})$, independently for each demographic subgroup. The solid lines, labeled "Actual," in each panel of Figure 8 plot the resulting series and confirm a well-known finding that the the rate of return to years worked increase more steeply for college than high school educate workers, and for women relative to men in the 1970s and 1980s.

Next, we ask whether the estimates in Section 4 of our aggregate model with one state variable - the E/L ratio - can match the dynamics of the rate of return for each of the demographic

[21]We obtain similar results when considering the average rate of return to years worked across workers instead of the return to prior work of a "representative worker."



(a) By Education. (b) By Gender.

Figure 8: Actual and Predicted Rates of Return to Years Worked for the Representative Worker from Different Demographic Subgroups and Counterfactuals with Constant Relative Price of Experience.

subgroups. Remarkably, the answer is affirmative, as is illustrated by the lines labeled "Predicted" in Figure 8. We view this as strong evidence that the heterogeneous dynamics across demographic groups is indeed driven by the evolution of *one* aggregate price. As the stocks of labor and experience differ across demographic subgroups, their earnings respond differently to the changes in the common aggregate relative price of these inputs. This is further confirmed by the results of a counterfactual experiment of imposing $\mu = 1$ to eliminate the effects of the changes in the aggregate supply of experience on the relative price of experience (plotted in the line labeled "Composition").

This evidence also allows us to assess the role of selection in driving our results. It is theoretically possible that the increased labor force participation of women (where a non-random subset of women decide to join the labor force) or a selection of who goes to college affected the dynamics of Π and E/L, potentially inducing a correlation between them and biasing our estimate of μ. The results in this section imply that this does not appear to be the case. Indeed, if the co-movement of the aggregate Π and E/L was induced by such selection, we would not expect to fit the separate dynamics of $\Omega\left(s_{it}, x_{it}\right)$ for each of the subgroups.

5.4 Accounting for the Changing Cohort-Based Life-Cycle Profiles of Earnings

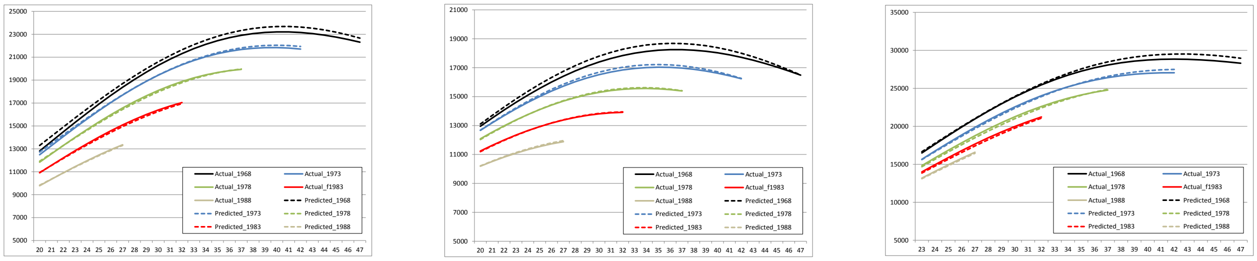
As we discussed in the Introduction, cross-sectional and cohort-based life-cycle earnings profiles diverge when the price of experience changes over time. Kambourov and Manovskii (2005, 2009) have noted that a substantial steepening of the cross-sectional profiles in the 1970s and 1980s (which was the primary focus of our analysis so far) was accompanied by a flattening of life-cycle earnings profiles for successive cohorts of male workers entering the labor market in the 1970s and 1980s. This means, for example, that a member of a large cohort of workers entering the labor market in, say, the 1970s, earned a low wage relative to older workers at that time. However, this does not imply that the wage growth over the life-cycle of this individual was going to be high. On the contrary, it was relatively low as well.

The success of our model in matching the dynamics of cross-sectional earnings distributions translates into its ability to account for the changes in the cohort-based life-cycle profiles.[22] Here, we illustrate its performance by its ability to account for the flattening of cohort based profiles of male workers documented in Kambourov and Manovskii (2005). We start by replicating the analysis in that paper which uses the PSID data over the 1968-1997 period on male workers employed full-time full-year (we use a cutoff of 1800 hours per year). On that sample, the following regression model is estimated:

$$y_{it} = \beta_0 + \beta_1 z_i + \beta_2 z_i^2 + \beta_3 z_i x_{it} + \beta_4 x_{it} + \beta_5 x_{it}^2 + \beta_6 x_{it}^3 + \epsilon_{it}, \tag{29}$$

where y_{it} is the log average real annual earnings of cohort i in period t, x_{it} is the age of cohort i in period t, z_i is the entry year of cohort i, and ϵ_{it} is a white noise term. The quadratic in the cohort entry year allows for different profile intercepts for different cohorts. The cubic in age gives all cohorts a similar shape, while the interaction of the linear age and cohort terms allows different cohorts to have different slopes of the earnings profiles. For instance, if the coefficient on the interaction term is negative, then every successive cohort has a flatter earnings profile. This regression uses information from all the cohorts present in the labor market and provides a very good fit to the data. The fitted profiles for selected cohorts are plotted in the lines labeled

[22] In Appendix A9, we derive some results useful for understanding the effect of cohort size on its earnings.



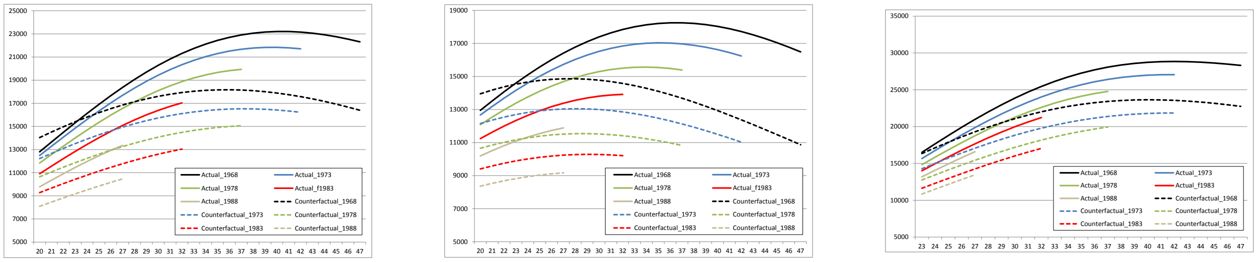
(a) All Education Levels. (b) High School. (c) College.

Figure 9: Actual and Predicted Life-Cycle Earnings Profiles.

"Actual" in Figures 9 and 10. The analysis is performed separately on the sample of high school educated men, college educated men, and on the combined sample.

Next we apply the same procedure to fitted values implied by the structurally estimated parameters of our model, including the technology parameters $\hat{\mu}$, $\hat{\delta}$, and the changing relative supply of experience.[23] The resulting predicted cohort-based life-cycle profiles are plotted in the lines labeled "Predicted" in Figure 9. The fit is very good implying that our model captures nearly perfectly the changes in cohort based profiles (the associated coefficients β_3 on the age-cohort interaction term governing the change in the profiles are reported in Table 3). If we restrict labor and experience to be perfect substitutes by setting $\mu = 1$, the model loses its ability to match the pattern of changes in cohort based profiles as can be seen in the lines labeled "Counterfactual" in Figure 10 and associated coefficients in Table 3. Thus, we conclude that the changes in the cohort-based life-cycle profiles of earnings are driven by the dynamics of the price of experience, which is determined, in turn, by the evolution of the relative supply of experience.

[23] The model is estimated on our standard sample in this paper, which is longer in time span, includes men and women, and does not impose the full-time full-year restriction. Then, we select fitted wages for the subsample satisfying the restrictions in Kambourov and Manovskii (2005). On this subsample of fitted values the model in equation (29) is estimated with the resulting predicted life-cycle profiles plotted in Figure 9. The same procedure is used to construct the "Counterfactual" predicted profiles in Figure 10.



(a) All Education Levels. (b) High School. (c) College.

Figure 10: Actual Life-Cycle Earnings Profiles and Counterfactuals with Constant Relative Price of Experience.

Table 3: Estimates of the Coefficient β_3 on Age-Cohort Interaction.

Specification	All Education Levels (1)	High School (2)	College (3)
Data	-.0003536	-.0005936	-.0001294
	(.0001237)	(.0001535)	(.0002046)
Predicted	-.0002742	-.0006262	-.0000722
	(.0000878)	(.0000917)	(.0000814)
Counterfactual	.0005318	.0002086	.0006231
	(.0001074)	(.0001101)	(.0000946)

Note - Standard errors in parentheses.

5.5 Aggregate Productivity

Given the estimates of $\widehat{\mu}$ and $\widehat{\delta}$, we can uncover the marginal product of labor G_{L_t} implied by the specification of the aggregate production function in (8):

$$\widehat{G}_{L_t} = \left(1 + \widehat{\delta}\left(\frac{\widehat{E}_t}{\widehat{L}_t}\right)^{\widehat{\mu}}\right)^{\frac{1}{\widehat{\mu}}-1}. \tag{30}$$

Combining $\widehat{G}_{L_t}$ and the estimates of the time-varying intercept terms $\widehat{D}_t$, the log of the aggregate productivity term (for labor earnings; see Footnote 4) A_t can be identified by

$$\ln \widehat{A}_t = \widehat{D}_t - \ln \widehat{G}_{L_t}. \tag{31}$$

Thus, we can decompose the changes in D_t into a component due to changes in the experience-labor ratio $\frac{E_t}{L_t}$, and a component due to changes in aggregate productivity level A_t. The results

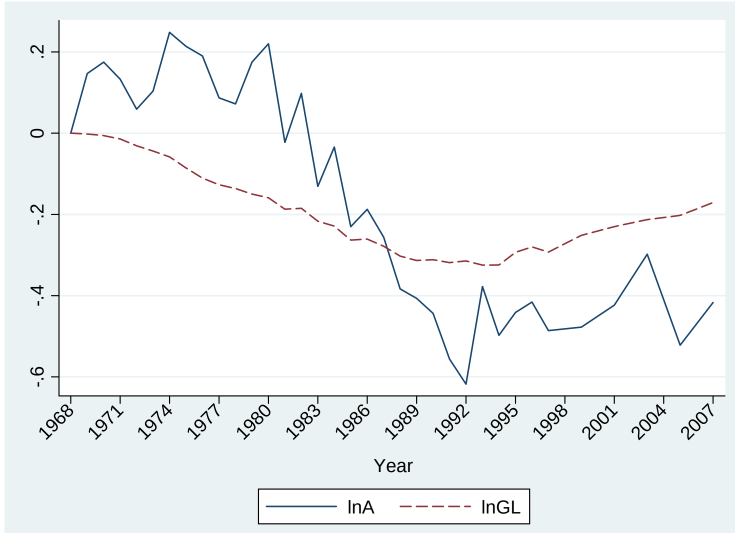


Figure 11: Aggregate Labor Productivity and Marginal Product of Labor.

of this decomposition are presented in Figure 11.

To facilitate comparisons of the movements among these variables, we normalize values in 1968 to zero (by subtracting a constant). The log of marginal product of labor $\ln \widehat{G}_{L_t}$ has decreased over the sample period accounting for most of the slowdown in the growth of D_t and its eventual decline. This translates into a substantial 26% fall in the level of the marginal product of labor between 1968 and 1996. Thus, the model not only accounts for the dynamics of the return to prior work, but also endogenously generates a substantial decline in the intercept of the wage equation. It is not a priori clear that these two features of the data might be related, but the model provides a tight link between them. When the experience to labor ratio declines, the marginal product of labor declines as well. This is exactly what the intercept of the wage equation captures.

6 Conclusion

In this paper we evaluated the role of the changing supply of experience, driven largely by the progression of the baby boom cohorts through the labor market, and by the increase in female labor force participation, in accounting for some of the key labor market trends over the last forty years. The main tool of our investigation is the aggregate production function that allows for complementarity between labor and experience - the two productive factors supplied by the workers to the labor market. We found that the evolution of the supply of experience accounts nearly perfectly for the large changes in the relative price of experience over the last forty years. It also accounts well for the changes in age premiums across education groups and changes in college premium across age groups. Finally, it accounts for the changes in the slopes and intercepts of cross-sectional and cohort-based life-cycle earnings profiles. While these developments were studied in separate strands of the literature, we find that the changing supply of experience provides a powerful unifying explanation.

Methodologically, our approach based on decompositions of individual earnings also allows us to relax some of the assumptions underlying the existing measurement approaches. In particular, we do not need to assume that some demographic groups exclusively supply "pure" labor or experience inputs. Instead, all workers can be supplying a bundle of these inputs, and we can measure the individual components of such bundles through variation in years actually worked by individuals of a given age. Consistent aggregation of individual earnings equations ensures that these quantities correspond to the objects implied by the aggregate production function. Moreover, our approach provides a way to filter out the effects of changes in, e.g., college or gender premium, when studying the evolution of the relative price of experience. The additional flexibility afforded by our approach appears important. In particular, while the existing literature ascribes a prominent role to changes in relative demand in shaping the empirical patterns we study, we assign a virtually exclusive role to changes in relative supply in accounting for them. This is not a forgone conclusion in our approach as it is not designed to favor either demand or supply based explanations.

At a more micro level, we find that separating the effects of age and experience in individual earnings provides a simple way to introduce cohort effects into traditional Mincerian wage equations. See Appendix A5 for a discussion of how our model generates age, time, and cohort effects. Interestingly, we find that the concavity of the age-wage profile is due to age effects rather than decreasing returns in experience input production.

The paper also offers new insights to the literature exploring the relationship between the size of a cohort and the relative earnings of its members. Motivated by the baby boom generation experience, Freeman (1979), Welch (1979) and Berger (1985) provided early empirical evidence that larger cohorts suffer depressed earnings upon entry into the labor market. More recent evidence is summarized in Wasmer (2001b,a) and Triest, Sapozhnikov, and Sass (2006). Jeong and Kim (2006) found related evidence in a dual economy model of transition for Thailand, while Kim and Topel (1995) found that a sharp decline in the share of young workers in South Korea was associated with an increase in their relative earnings. Despite this suggestive evidence, Topel (1997) summarizes this literature by saying: "The effects of cohort size on earnings tend to be a sideline in the inequality literature." Our theory and quantitative results (see also Appendix A9) show that demographic change arising from changes in cohort size may be key to understanding the dynamics of the returns to experience and the associated wage inequality across cohorts. In our setting, different cohorts do not supply distinct labor inputs, yet cohorts separated more in time will appear to complement each other because of their compositional difference in terms of labor and experience.

References

Berger, M. C. (1985): "The Effect of Cohort Size on Earnings Growth: A Reexamination of the Evidence," *Journal of Political Economy*, 93(3), 561–573.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 116(2), 705–746.

Carneiro, P., and S. Lee (2011): "Trends in Quality-Adjusted Skill Premia in the United States, 1960-2000," *American Economic Review*, 101(6), 2309–49.

Elsby, M. W. L., and M. D. Shapiro (2012): "Why Does Trend Growth Affect Equilibrium Employment? A New Explanation of an Old Puzzle," *American Economic Review*, 102(4), 1378–1413.

Freeman, R. B. (1979): "The Effect of Demographic Factors on Age-Earnings Profiles," *Journal of Human Resources*, 14(3), 289–318.

Guvenen, F., and B. Kuruscu (2009): "A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970-2000," *NBER Macroeconomics Annual*, 24, 231–276.

——— (2010): "Understanding the Evolution of the U.S. Wage Distribution: A Theoretical Analysis," *Journal of the European Economic Association*, 10(3), 482–517.

Heckman, J. J., L. Lochner, and C. Taber (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1, 1–58.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, and F. Welch, vol. 1, chap. 7, pp. 307–458. Elsevier.

Heckman, J. J., and J. A. Scheinkman (1987): "The Importance of Bundling in a Gorman-Lancaster Model of Earnings," *Review of Economic Studies*, 54(2), 243–55.

Jeong, H., and Y. Kim (2006): “Complementarity and Transition to Modern Economic Growth,” Working Paper 06.44, Institute for Economic Policy Research.

Kambourov, G., and I. Manovskii (2005): “Accounting for the Changing Life-Cycle Profile of Earnings,” 2005 Meeting Papers 231, Society for Economic Dynamics.

——— (2009): “Occupational Mobility and Wage Inequality,” *Review of Economic Studies*, 76(2), 731–759.

Katz, L. F., and K. M. Murphy (1992): “Changes in Relative Wages, 1963-87: Supply and Demand Factors,” *Quarterly Journal of Economics*, 107(1), 35–78.

Kim, D.-I., and R. H. Topel (1995): “Labor Markets and Economic Growth: Lessons from Korea’s Industrialization, 1970-1990,” in *Differences and Changes in Wage Structures*, ed. by R. B. Freeman, and L. F. Katz, pp. 227–64. Chicago: University of Chicago Press for NBER.

Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000): “Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis,” *Econometrica*, 68(5), 1029–1053.

Lemieux, T. (2006): “The Mincer Equation Thirty Years after Schooling, Experience, and Earnings,” in *Jacob Mincer, A Pioneer of Modern Labor Economics*, ed. by S. Grossbard, chap. 11, pp. 127–148. Ney York: Springer.

Light, A., and M. Ureta (1995): “Early-Career Work Experience and Gender Wage Differentials,” *Journal of Labor Economics*, 13(1), 121–154.

Mincer, J. (1958): “Investment in Human Capital and Personal Income Distribution,” *Journal of Political Economy*, 66(4), 281–302.

——— (1974): *Schooling, Experience, and Earnings*. Columbia University Press, New York, NY.

Song, Z. M., and D. T. Yang (2012): “Life Cycle Earnings and Saving in a Fast-Growing Economy,” mimeo, University of Chicago Booth School of Business.

Topel, R. H. (1997): “Factor Proportions and Relative Wages: The Supply-Side Determinants of Wage Inequality,” *Journal of Economic Perspectives*, 11(2), 55–74.

Triest, R. K., M. Sapozhnikov, and S. A. Sass (2006): "Population Aging and the Structure of Wages," CRR WP 2006-5, Center for Retirement Research at Boston College.

Tukey, J. W. (1977): *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.

Vuong, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57(2), 307–333.

Wasmer, E. (2001a): "Between-group Competition in the Labor Market and the Rising Returns to Skill: US and France 1964-2000," Discussion Paper No. 292, IZA.

——— (2001b): "Measuring Human Capital in the Labor Market: The Supply of Experience in 8 OECD Countries," *European Economic Review, Papers & Proceedings of the 15th Annual Congress of the European Economic Association*, 45(4-6), 861–874.

Weinberg, B. A. (2005): "Experience and Technology Adoption," Working Paper, Ohio State University.

Welch, F. (1979): "Effects of Cohort Size on Earnings: The Baby Boom Babies Financial Bust," *Journal of Political Economy*, 87(5), S65–S97.

A1 The Role of Age and Years of Prior Work: A Simple Two-by-Two Example

Consider an economy with 4 types of workers each earning $w_{j,t}^s$ in period t, where $s \in \{h, c\}$, h for high-school, c for college, and $j \in \{y, o\}$, y for young and o for old. Earnings of each group can be decomposed into a linear combination of payments for the two inputs – labor, L, and experience, E:

$$\begin{aligned} w_{j,t}^c &= c_t \left[a_j^c R_{L,t} + b_j^c R_{E,t} \right], \\ w_{j,t}^h &= a_j^h R_{L,t} + b_j^h R_{E,t}, \end{aligned}$$

where a_j^s denotes the time invariant quantity of input L supplied by schooling group s of age j, b_j^s denotes the corresponding supply of input E, c_t is a time varying aggregate shock to the productivity of college-educated workers, $R_{L,t}$ and $R_{E,t}$ denote the time varying economy wide prices of inputs L_t and E_t. Let $\Pi_t = R_{E,t}/R_{L,t}$ denote the relative price of the two inputs.

As in Katz and Murphy (1992), assume that all young workers exclusively supply one unit of L_t only ($a_y^c = a_y^h = 1, b_y^c = b_y^h = 0$). Old workers can potentially supply a combination of both inputs L_t and E_t, with the weights $\{a_o^c, a_o^h, b_o^h\}$ to be determined after normalizing $b_o^c = 1$. Decomposing log wages, we have

$$\begin{aligned} \ln w_{y,t}^c &= \ln c_t + \ln R_{L,t}, \\ \ln w_{o,t}^c &= \ln c_t + \ln R_{L,t} + \ln \left[a_o^c + \Pi_t \right], \\ \ln w_{y,t}^h &= \ln R_{L,t}, \\ \ln w_{o,t}^h &= \ln R_{L,t} + \ln \left[a_o^h + \Pi_t b_o^h \right]. \end{aligned}$$

The first and third equations imply that $\ln R_{L,t}$, $\ln c_t$ can be readily identified by the variation in $\ln w_{y,t}^h$ and $\ln w_{y,t}^c - \ln w_{y,t}^h$. The remaining parameters to be identified are a_o^c, a_o^h, b_o^h, and Π_t. It is the identification strategy for these parameters where our approach differs from that of Katz and Murphy.

A1.1 Katz and Murphy's Identification Strategy

Katz and Murphy assume that old workers completely stop supplying the input that young workers supply and simply set $a_o^c = a_o^h = 0$. This resolves the identification problem as the age premium for college and high school educated workers is respectively given by

$$\frac{w_{o,t}^c}{w_{y,t}^c} = \Pi_t \text{ and } \frac{w_{o,t}^h}{w_{y,t}^h} = \Pi_t b_o^h.$$

One consequence of this identifying assumption, evident from the equation above, is that the elasticity of the age premium with respect to Π_t is forced to be the same across college and high school educated workers (while these age premiums move differently in the data). As the ratio of age premiums between college and high school workers is equal to the ratio of college premiums between old an young workers, this identifying assumption also restricts the potential role of changes in Π_t in explaining the differential movement of college premiums across age groups that motivated the analysis in Card and Lemieux (2001).

A1.2 Our Identification Strategy

Our identification strategy is based on the idea that the relevant parameters can be identified if old workers of the same age differ in the number of years they have actually worked (and consequently accumulated different amounts of the experience input). To use an extreme example, note that we can identify a_o^c and Π_t if we observe some old college workers who never worked and thus accumulated no experience input. They earn $\tilde{w}_{o,t}^c = c_t R_{L,t} a_o^c$ in contrast to old college workers whose years of prior work endowed them with one unit of experience input and who earn $w_{o,t}^c = c_t \left[R_{L,t} a_o^c + R_{E,t}\right]$. Similarly, we can identify a_o^h and $\frac{b_o^h}{a_o^h}$ if we observe some old high school workers who accumulated no experience input and earn $\tilde{w}_{o,t}^h = R_{L,t} a_o^h$ and compare them with old high school workers whose years of prior work endowed them with b_o^h units of experience input so that they earn $w_{o,t}^h = R_{L,t} a_o^h + R_{E,t} b_o^h$. Using the PSID data we are able to effectively do this from the imperfect correlation between age and the number of years that individuals actually worked by that age. While the two-age example here is extreme, we show in Appendix A3 that a small amount of variation in the number of years worked is sufficient to fully identify the model in the main text non-parametrically.

Because we do not impose the assumption that old workers exclusively supply the experience input, the age premiums across schooling groups are given by

$$\frac{w^c_{o,t}}{w^c_{y,t}} = a^c_o + \Pi_t \text{ and } \frac{w^h_{o,t}}{w^h_{y,t}} = a^h_o + \Pi_t b^h_o,$$

which implies the elasticity of the age premium with respect to Π_t differs across schooling groups depending on $\frac{1}{a^c_o} \lesseqgtr \frac{b^h_o}{a^h_o}$. Our estimates of these parameters in the main text allow us to account for the different movements of the age premiums across schooling groups and college premiums across age groups.

A2 PSID Data

Sample. We use the Panel Study of Income Dynamics (PSID) data from the U.S. for the 1968-2007 period. The PSID consists of two main subsamples: the SEO (Survey of Economic Opportunity) sample and the SRC (Survey Research Center) sample. We use both samples and restrict ourselves to the core members with positive sampling weights (not the newly added family members through marriage) to maintain the consistent representativeness of the sample over time.[24] The sample is restricted to individuals between 18 and 65 years of age.

Years of Prior Work. The procedure we use to construct measures of actual years worked since age 18 is as follows. Questions regarding the number of years worked ("How many years have you worked for money since you were 18?" and "How many of these years did you work full time for most or all of the year?") were asked of every household's head and wife in 1974, 1975, 1976 and 1985.[25] These questions are also asked for every person in the year when that person first becomes a household head or wife.[26] Since there are some inconsistencies between the answers, we first adjust the 1974 report to be consistent with 1975 and 1976 values when

[24] We use only the nonimmigrant sample. In 1990 the PSID added a new sample of 2000 Latino households, consisting of families originally from Mexico, Puerto Rico, and Cuba. Because this sample missed immigrants from other countries, Asians in particular, and because of a lack of funding, this Latino sample was dropped after 1995. Another sample of 441 immigrant families was added in 1997. Because of the inconsistencies in these samples, we restrict ourselves to the core SEO and SRC samples throughout the 1968-2007 period.

[25] By default, the head of household is the (male) husband if he is present or a female if she is single. In very few cases the head is a female, even when the male husband is present (but is, say, severely disabled).

[26] The PSID mistakenly did not ask some people in 1985 and fixed this mistake by asking them in 1987.

possible. Next, we use 1974 as the base year; i.e., we assume that whatever is recorded in 1974 for the existing heads is true. For the entrants into the sample we assume that the number of years of prior work they report in their first year in the sample is true. If the report implies that an individual started working before the age of 18, we redefine it to be the number of years since age 18 for that individual. If the reported number of years worked in 1974 is smaller than that implied by the reports of hours between the individual entry into the sample (or 1968) and 1974, we replace the 1974 report with that implied by the accumulated reports of hours. We repeat this procedure for 1985 and for the reports of the new heads and wives. Finally, using the values in 1974, 1985, and the reports of the new heads and wives, we increment the years of work variables forward and backward as follows: increment the full-time measure by one if individual works at least 1500 hours in a given year.[27] If we observe an individual in the sample since age 18, we ignore his or her reports and instead directly use his or her reports of hours in each year using the cutoff above.[28]

Other Variables. Our hourly wage measure is equal to the total earnings last year divided by total hours worked last year. To get the real wage, we adjust the nominal wage using last year's CPI (equal to 100 in 1984).[29] We define the economically active population as the group of people who worked at least 700 hours last year.[30] Education is measured by years of final educational attainment.[31] Other control variables that we will use are gender (male dummy),

[27]We experimented with using cutoff values other than 1500 hours of work or using directly the sum of accumulated hours of work to create other measures of prior work and found that our chosen measure shows the smoothest pattern of movements. The substantive results are not sensitive to this choice.

[28]The PSID switched from annual to bi-annual interviewing after 1997. Some data for the non-interview years is available but appears very noisy with large numbers of missing observations. This led us to use only the data from years when interviews took place. The only exception is hours worked in years between interviews which are needed to construct the measures of prior work. We imputed those hours as the maximum between the reported hours (if available) and the average hours in the two adjacent survey years.

[29]There is an alternative hourly wage measure in the PSID which reports the current hourly wage at the time of the interview. Unfortunately, this measure is only available for the household heads throughout the period. For wives it is available only in 1976 and after 1979 and it is not available at all for the other family members.

[30]As in the case of earnings, there is also an employment status variable at the time of the interview. We do not use this variable because (1) the reference period (current year) is different from that of the earnings measure (last year), and (2) this variable is available for the heads for all years but not for the wives before 1979 except in 1976 and is not available for the dependents.

[31]Education is reported in the PSID in 1968, 1975, and 1985 for existing heads of households, and every year

race (black dummy), and region (Northeast, North Central and West dummies). The broad region variable is created using the state variable in the PSID.[32] South is the base category region.

A3 Variation in Years Worked Needed for Identification

As is well known, the variation in the number of years worked by a certain age is relatively small, especially for male workers. This might appear to pose a challenge for our identification strategy. We now show that the model is nonparametrically identified if we only have one year of difference in the number of years worked at each age, e.g., it is enough that some workers enter the labor market at age 18 while some others at age 19 so that at age 20, they have worked 2 and 1 years, respectively, and at age 21, they have worked 3 and 2 years, respectively, etc. Following this, we show that there is much more variation available in our data.

A3.1 Non-parametric Identification

We now establish non-parametric identification of the relative price of experience Π_t and of the $\lambda_L(j)$, $\lambda_E(j)$ and $g(e)$ schedules if within each age group j some individuals worked for j years and some others for $j-1$ years, i.e. $e \in \{j, j-1\}$ for $j \geq 1$.

Specify the log wage of a worker with age j and years worked e as

$$\ln w(j,e) = \ln R_{L,t} + \ln(\lambda_L(j) + \Pi_t \lambda_E(j) g(e)).$$

for the people becoming household heads or wives. It is kept constant between the years in which it is updated. As a result, there would be a bias toward a lower educational level. For example, if education is 10 years in 1975 and 16 in 1985, it would be reported 10 between 1975 and 1985. If the individual, however, had 16 years of education already in 1980, then for five years he would be counted as less educated than he actually is. To minimize this bias, the education variable used in the estimation is fixed to be equal to its mode value among all the reports available. To make the education variable comparable across time we top code it at 16 years.

[32] We found that the broad region variable provided by the PSID appears to be error-ridden. For example, for some but not all Texas residents region is defined as West. Thus, we reconstructed the broad region variable directly from the state of residence.

We have the restriction that $\lambda_L(0) = 1$ and $g(0) = 0$, so that

$$\begin{aligned}\ln w(0,0) &= \ln R_{L,t},\\ \ln w(1,0) &= \ln R_{L,t} + \ln \lambda_L(1),\end{aligned}$$

which are used to identify $\ln R_{L,t}$ for all t and $\ln \lambda_L(1)$. Now consider

$$\begin{aligned}\ln w(j, e=j) &= \ln R_{L,t} + \ln(\lambda_L(j) + \Pi_t \lambda_E(j) g(e=j))\\ \ln w(j, e=j-1) &= \ln R_{L,t} + \ln(\lambda_L(j) + \Pi_t \lambda_E(j) g(e=j-1))\end{aligned}$$

for all $j \geq 1$ and for every t. Since $\ln R_{L,t}$ is known, these equations imply that we can determine the following after exponentiating

$$\begin{aligned}&\lambda_L(j) + \Pi_t \lambda_E(j) g(e=j), \qquad \text{(A1)}\\ &\lambda_L(j) + \Pi_t \lambda_E(j) g(e=j-1),\end{aligned}$$

for all $j \geq 1$ and for every t. Since $\ln \lambda_L(1)$ is also known, we can further determine

$$\Pi_t \lambda_E(1) g(1)$$

for every t, from which we can determine $\frac{\Pi_{t+1}}{\Pi_t}$ for every t.

Using (A1) and the time differences of these values, we can find

$$\begin{aligned}&(\Pi_{t+1} - \Pi_t)\, \lambda_E(j) g(e=j),\\ &(\Pi_{t+1} - \Pi_t)\, \lambda_E(j+1) g(e=j),\\ &(\Pi_{t+1} - \Pi_t)\, \lambda_E(j+1) g(e=j+1).\end{aligned}$$

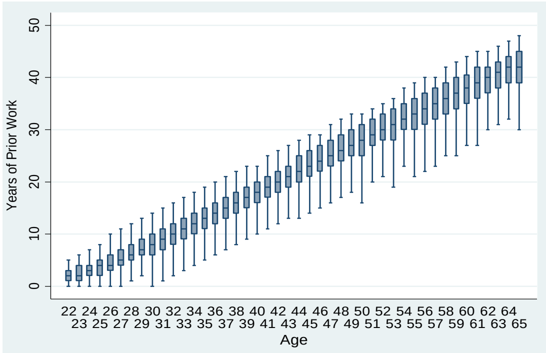
Using ratios of these, we can determine $\frac{\lambda_E(j+1)}{\lambda_E(j)}$, $\frac{g(e=j+1)}{g(e=j)}$ beginning from $\frac{\lambda_E(2)}{\lambda_E(1)}$, $\frac{g(2)}{g(1)}$.

Next, to determine $\lambda_L(j)$ we can use

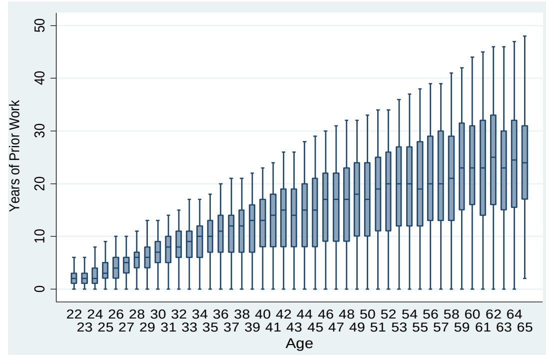
$$\frac{\lambda_L(j) + \Pi_t \lambda_E(j) g(e=j)}{\Pi_t \lambda_E(1) g(1)} = \frac{\lambda_L(j)}{\Pi_t \lambda_E(1) g(1)} + \frac{\lambda_E(j) g(e=j)}{\lambda_E(1) g(1)},$$

where the second term on the right hand side is known given the calculated $\frac{\lambda_E(j+1)}{\lambda_E(j)}$, and $\frac{g(e=j+1)}{g(e=j)}$, and $\Pi_t \lambda_E(1) g(1)$ is also known, allowing us to identify $\lambda_L(j)$.

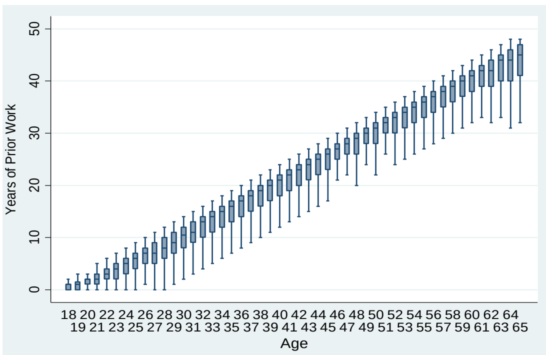
Finally, although from $\Pi_{t=0} \lambda_E(1) g(1)$ we cannot separately identify $\Pi_{t=0}$ and $\lambda_E(1)$ and $g(1)$, we could normalize two of these, $\lambda_E(1)$ and $g(1)$, without loss of generality to identify



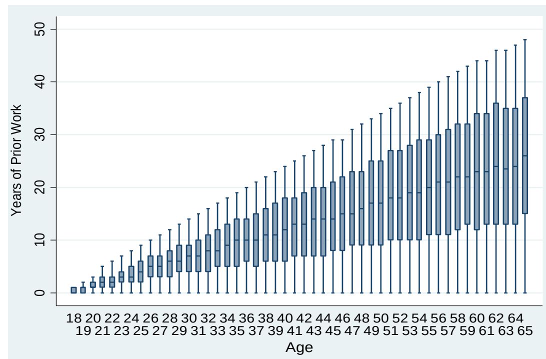
(a) College Male



(b) College Female



(c) High School Male



(d) High School Female

Figure A-1: Within-Age Variation in Years of Prior Work by Gender and Education.

$\Pi_{t=0}$.[33] As discussed in Appendix A7, the level of $\Pi_{t=0}$ does not affect our substantive results and it only scales the level of the estimated share parameter $\ln \delta$ in the aggregate technology (not the estimate of the complementarity parameter). In the specification used in the main text this normalization is not needed because $\lambda_E(0) = 1$, $g\,(0) = 0$ and the restricted functional forms of $\lambda_E(j)$ and $g\,(e)$.

Note that given the nonparametric identification achieved, the parametric identification in the main text is guaranteed as a special case.

A3.2 Variation in Years Worked Available in the Data

Figure A-1 uses boxplots to summarize the amount of variation in actual experience by age available in our data. Each plot shows the percentile statistics of the distribution such as median, 25th percentile, 75th percentile in a box and the upper and lower adjacent values in

[33]This is the same normalization we used when setting $b_o^c = 1$ in Appendix A1.

marking boundary values.[34] The figure illustrates that the range of variation of the number of years worked for every age group far exceeds the amount of variation needed for identification. Even among male workers, the effective range of within-age variation in the number of years worked is wider than 10 years for most age groups.

An alternative way to describe the amount of variation available for identification in the literature would be to to report the correlations between age and years of prior work. In our data these are 0.95 for college males, 0.95 for high-school males, 0.73 for college females, and 0.66 for high-school females. The interpretation of such correlations is, however, not straightforward in the context of establishing identification. This is because the overall correlation is dominated by the overall co-movement in age and years of prior work and does not immediately reveal the extent of years of prior work variation conditional on age, which determines the identification. The following simple example illustrates this.

Suppose that for each age the distribution of the number of years worked is constant around that age (the sample size N also ensures that this is the case). Let x_i denote age and y_i the years of prior work, then

$$y_i = x_i + e_i$$

where $e_i = e$ is from a given distribution independent of the level of x_i such that

$$\sum_i [e_i (x_i - \bar{x})] = 0.$$

Without loss of generality, re-normalize the measure of years worked such that the average number of years worked and age are equal, $\bar{x} = \bar{y}$.

[34] The "upper and lower adjacent values" are the extreme values of ± 1.5 times of the inter-quartile range, which are suggested by Tukey (1977) to capture the "effective range" of the distribution.

The aggregate correlation between age and years of prior work is given by

$$\begin{aligned}
r &= \frac{\sum_i \left[(x_i - \bar{x})(y_i - \bar{y})\right]}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \\
&= \frac{\sum_i \left[(x_i - \bar{x})(x_i - \bar{x} + e_i)\right]}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (x_i - \bar{x} + e_i)^2}} \\
&= \frac{\sum_i (x_i - \bar{x})^2}{\sqrt{\sum_i (x_i - \bar{x})^2 \left[\sum_i (x_i - \bar{x})^2 + \sum_i (e_i)^2\right]}} \\
&= \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{\sum_i (x_i - \bar{x})^2 + \sum_i (e_i)^2}}.
\end{aligned}$$

Thus, the correlation is falling in the ratio

$$\frac{\sum_i (e_i)^2}{\sum_i (x_i - \bar{x})^2} = \frac{N\sigma_e^2}{\sum_i^N (x_i - \bar{x})^2}.$$

Since the numerator is constant, this ratio is essentially falling in the range of ages in the sample, whereas the variation of age and years of prior work that is relevant for the identification is given by σ_e^2.

To provide a quantitative example, if we partition our sample into 10 equally spaced birth cohort bins, the average correlation within a bin is 0.9 for college males, 0.9 for high-school males, 0.58 for college females, and 0.57 for high-school females. To avoid arbitrariness of choosing such a partition, we think Figure A-1 is more informative in summarizing the ample variation available for identification in our data.

A4 Descriptive Analysis, Additional Results

A4.1 Descriptive Analysis, Benchmark Coefficient Estimates

Table A-1: Descriptive Analysis, Estimates of Time-Invariant Parameters.

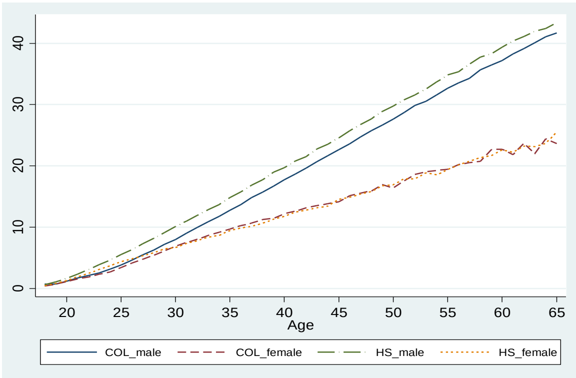
Parameter	Estimate	Standard error	t-statistic
$\lambda_{L,1}\,(HS, M)$	.0245	.00323	7.56
$\lambda_{L,2}\,(HS, M)$	-.000492	.0000803	-6.13
$\lambda_{L,0}\,(C, M)$	-.306	.0350	-8.74
$\lambda_{L,1}\,(C, M)$	.0671	.00363	18.47
$\lambda_{L,2}\,(C, M)$	-.00116	.0000846	-13.66
$\lambda_{E/L,1}\,(HS, M)$	-.0792	.00829	-9.55
$\lambda_{E/L,2}\,(HS, M)$	.00107	.000181	5.90
$\lambda_{E/L,0}\,(C, M)$	.332	.121	2.74
$\lambda_{E/L,1}\,(C, M)$	-0.148	.0132	-11.21
$\lambda_{E/L,2}\,(C, M)$	.00214	.000268	8.00
$\lambda_{L,1}\,(HS, F)$	.00109	.00203	.54
$\lambda_{L,2}\,(HS, F)$	.0000709	.0000463	1.53
$\lambda_{L,0}\,(C, F)$	-.0569	.0282	-2.02
$\lambda_{L,1}\,(C, F)$	.0345	.00265	12.98
$\lambda_{L,2}\,(C, F)$	-.000571	.0000612	-9.33
$\lambda_{E/L,1}\,(HS, F)$	-.0434	.00661	-6.57
$\lambda_{E/L,2}\,(HS, F)$	.0000404	.000137	.30
$\lambda_{E/L,0}\,(C, F)$	-.499	.127	-3.93
$\lambda_{E/L,1}\,(C, F)$	-.054	.0129	-4.23
$\lambda_{E/L,2}\,(C, F)$	.000142	.0000286	.50
θ_1	-.0234	.00476	-4.92
θ_2	.000979	.000184	5.32
θ_3	-.0000141	.00000238	-5.93
northeast	.19	.004	43.64
north central	.046	.004	11.46
west	.098	.0045	21.90
R^2	0.924		
$RMSE$	0.616		

Note - The entries represent the results of the reduced-form estimation of time-invariant parameters of the benchmark specification in Section 3. For sample restrictions and variable construction procedures, see Appendix A2.

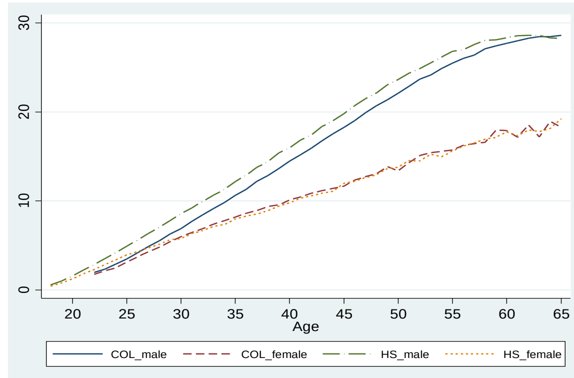
Table A-2: Descriptive Analysis, Estimates of Time-Varying Parameters.

Year	Price of Exp.	Schooling	Male	Black	Intercept
1968	.091(.015)	.072(.005)	.24(.031)	-.16(.048)	.46(.09)
1969	.078(.013)	.066(.005)	.25(.028)	-.16(.044)	.61(.08)
1970	.089(.013)	.063(.005)	.24(.027)	-.14(.042)	.63(.08)
1971	.087(.013)	.069(.005)	.22(.027)	-.10(.042)	.58(.08)
1972	.106(.014)	.074(.005)	.20(.026)	-.12(.041)	.49(.08)
1973	.104(.013)	.072(.004)	.25(.026)	-.086(.040)	.52(.07)
1974	.101(.012)	.064(.004)	.23(.023)	-.099(.036)	.65(.07)
1975	.109(.012)	.067(.004)	.21(.023)	-.087(.035)	.59(.07)
1976	.126(.013)	.069(.004)	.16(.023)	-.051(.035)	.54(.07)
1977	.131(.014)	.076(.004)	.19(.023)	-.037(.035)	.42(.07)
1978	.152(.015)	.075(.004)	.19(.023)	-.051(.035)	.40(.07)
1979	.141(.014)	.070(.004)	.20(.022)	-.087(.032)	.49(.07)
1980	.137(.013)	.066(.004)	.20(.021)	-.045(.031)	.53(.07)
1981	.170(.016)	.080(.004)	.17(.022)	-.092(.032)	.26(.07)
1982	.155(.015)	.071(.004)	.17(.022)	-.093(.032)	.38(.07)
1983	.206(.019)	.083(.004)	.10(.021)	-.075(.032)	.12(.07)
1984	.183(.016)	.081(.004)	.11(.020)	-.082(.030)	.20(.07)
1985	.217(.018)	.092(.004)	.15(.020)	-.078(.030)	-.03(.07)
1986	.207(.018)	.092(.004)	.12(.020)	-.12(.030)	.02(.07)
1987	.224(.015)	.099(.004)	.10(.020)	-.15(.029)	-.07(.07)
1988	.251(.021)	.108(.004)	.05(.020)	-.14(.029)	-.22(.07)
1989	.242(.019)	.111(.004)	.07(.018)	-.15(.026)	-.26(.07)
1990	.231(.019)	.114(.004)	.04(.018)	-.11(.026)	-.29(.07)
1991	.217(.018)	.123(.004)	.04(.018)	-.07(.027)	-.41(.07)
1992	.252(.020)	.122(.004)	.03(.018)	-.13(.026)	-.47(.07)
1993	.220(.018)	.111(.004)	.04(.018)	-.094(.026)	-.24(.07)
1994	.201(.017)	.119(.004)	.09(.018)	-.16(.026)	-.36(.07)
1995	.229(.019)	.110(.004)	.08(.018)	-.11(.026)	-.27(.07)
1996	.238(.020)	.106(.004)	.08(.018)	-.14(.026)	-.23(.08)
1997	.214(.018)	.115(.004)	.03(.017)	-.15(.025)	-.31(.07)
1999	.217(.018)	.113(.004)	.07(.017)	-.15(.025)	-.26(.07)
2001	.174(.016)	.117(.004)	.07(.017)	-.088(.024)	-.19(.07)
2003	.165(.015)	.106(.004)	.02(.016)	-.11(.024)	-.05(.07)
2005	.185(.016)	.120(.004)	.04(.016)	-.15(.024)	-.26(.07)
2007	.143(.014)	.118(.004)	.07(.016)	-.17(.023)	-.12(.07)

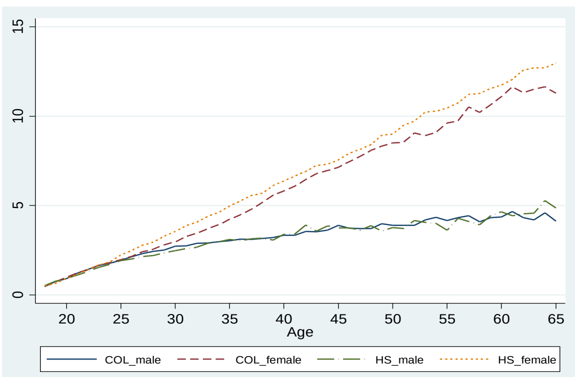
Note - The entries represent the results of the reduced-form estimation of time-varying parameters of the benchmark specification in Section 3. Standard errors are in parenthesis. For sample restrictions and variable construction procedures, see Appendix A2.



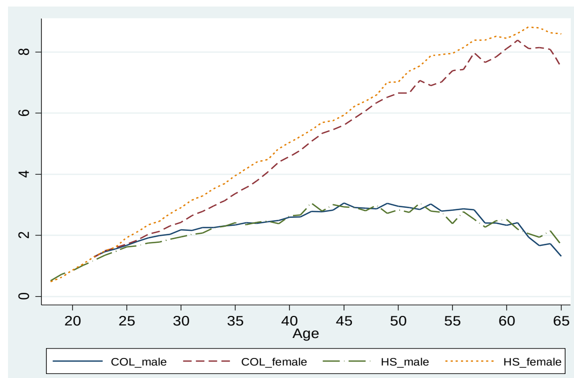
(a) Years of Prior Work, Mean



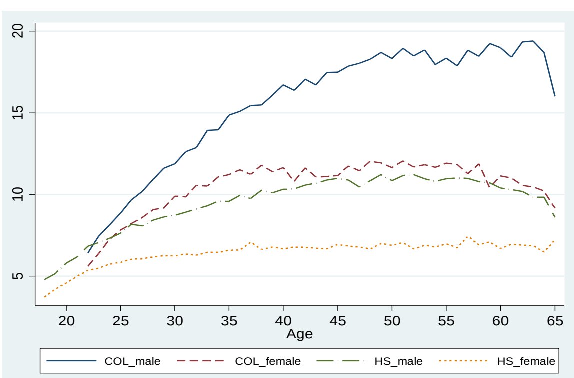
(b) Experience Input, $g(e)$, Mean



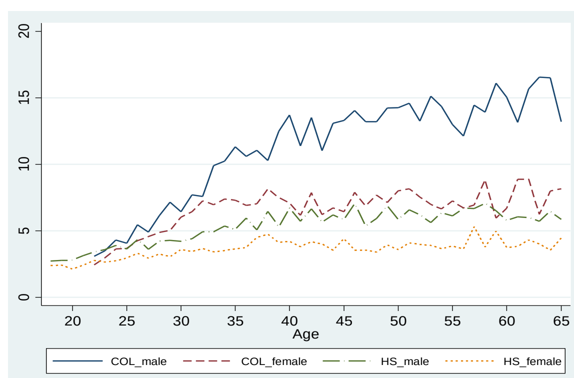
(c) Years of Prior Work, SD



(d) Experience Input, $g(e)$, SD

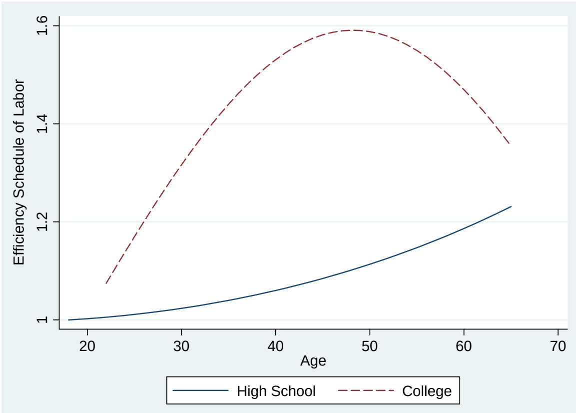


(e) Wage, Mean

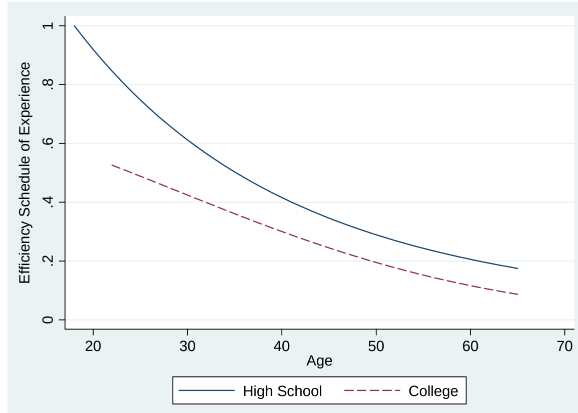


(f) Wage, SD

Figure A-2: Age-Conditional Means and Standard Deviations of Years Worked, Experience Input, and Wages by Gender and Education.



(a) Age Efficiency Schedule of Labor.



(b) Age Efficiency Schedule of Experience.

Figure A-3: Estimated Age-Efficiency Schedules for Female Workers by Education.

A5 Time-Invariant Age Efficiency Schedules and Cohort Effects

While our identification strategy follows Katz and Murphy (1992) and subsequent literature in assuming that age efficiency schedules λ_L and λ_E are independent of time, this is potentially an important restriction ruling out certain cohort effects. In this Appendix we empirically assess this assumption through two experiments. First, we check for the presence of cohort effects not accounted for by the model with time-invariant age efficiency schedules. Second, we estimate the model separately on different cohorts and check whether the estimates differ significantly.

A5.1 Cohort Effects

To check for the presence of cohort effects that are not accounted for by our specification with constant age efficiency schedules, we we obtain wage residuals from our model and ask whether we can detect the presence of residual cohort effects in them. In particular, we regress the residuals on the full set of cohort dummies. The estimates of coefficients on those dummies are all statistically insignificant from zero. Moreover, they do not exhibit any particular trends as is evident from Figure A-4 in which the coefficients of the cohort dummies are plotted for all cohorts.

How does our model account for cohort effects? The time effect on wages is captured via the changes in the year dummies of the wage regression (equation (16) in the main text). The pure

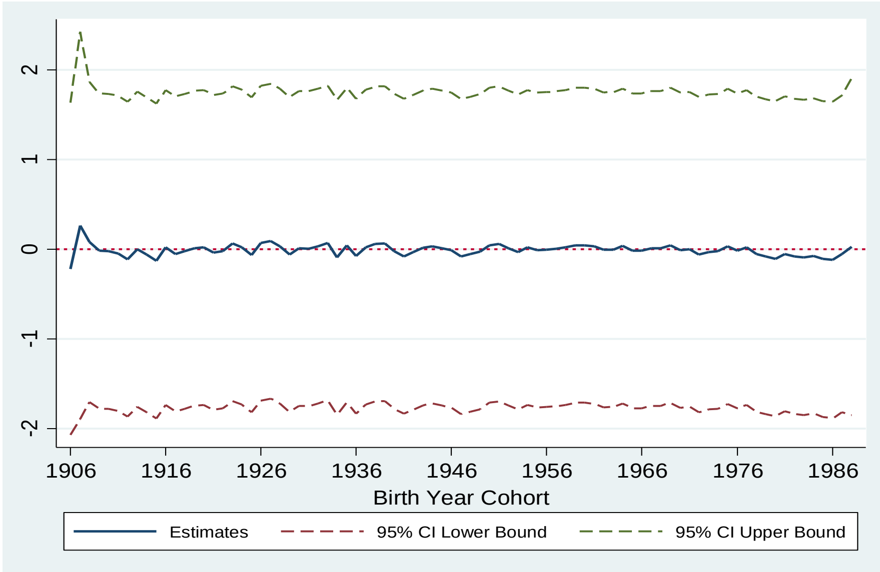


Figure A-4: Residual Cohort Dummy Estimates.

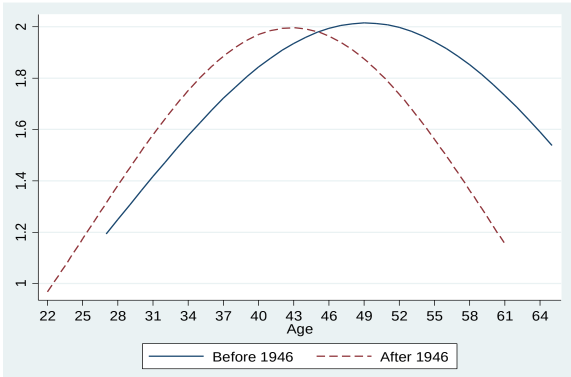
age effect is captured by the age efficiency schedule of labor $\lambda_L(j)$. The cohort effect is captured by the interactive term between the relative price of experience $\Pi_{E,t}$ (time effect) and the age efficiency schedule of labor $\lambda_{E/L}(j)$ (age effect).

Moreover, we also allow the time-varying coefficients for the characteristics of schooling, gender and race. Through the changing age composition of these demographic subgroups, allowing the time-varying coefficients on these characteristics also captures the cohort effect indirectly. These are clearly one particular way of capturing cohort effects and the question is whether we are missing other significant cohort effects than the ones captured. The results of the experiment in this section suggest that we are not.

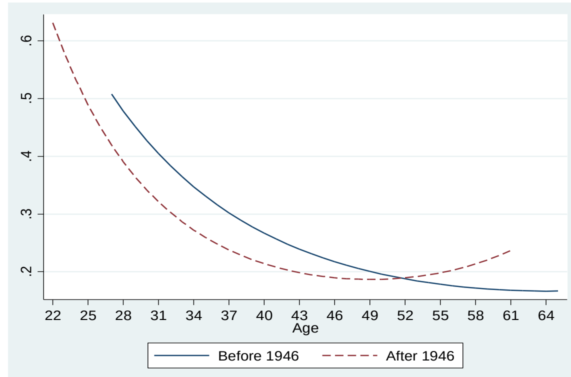
A5.2 Estimating the Model on Different Cohorts

We now allow the age efficiency schedules to differ across cohorts and check whether the estimates differ significantly. In particular, we allow the age efficiency schedules for males to be cohort specific as follows

$$
\begin{aligned}
\lambda_L^s(j, Z) &= \exp(\lambda_{L,1}^s j + \lambda_{L,2}^s j^2 + I_Z \left(Z_{L,1}^s j + Z_{L,2}^s j^2\right)), \\
\lambda_E^s(j, Z) &= \exp(\lambda_{E,1}^s j + \lambda_{E,2}^s j^2 + I_Z \left(Z_{E,1}^s j + Z_{E,2}^s j^2\right)),
\end{aligned}
$$



(a) λ_L



(b) λ_E

Figure A-5: Estimated $\lambda_L(j), \lambda_E(j)$ for Cohorts of College-Educated Male Workers Born before or after 1946.

where I_Z is an indicator for cohort Z.

Defining cohorts too finely runs into the problem of cohorts not having observations along the support of the age profile for the youngest and oldest cohorts, over and above sample size issues. Thus, we consider two cohort groups, those born before and those born after 1946. We choose the partition of cohorts at 1946 to generate the most overlap in terms of age between the two cohorts for comparison. This turns out to also imply similar sample sizes for the two groups.[35]

For the high school group, the four coefficients $\{Z^s_{L,1}, Z^s_{L,2}, Z^s_{E,1}, Z^s_{E,2}\}$ are all insignificant so there are no statistically relevant differences in λ^s_L and λ^s_E across cohorts. For the college group each of these four coefficients turned out to be statistically significant, and we investigated further the implications.

Plotting $\lambda^s_L(j, Z)$ and $\lambda^s_E(j, Z)$ over the support of age j for the pre-1946, post-1946 cohorts, in Figure A-5, we observe that the age efficiency schedules are quite similar in shape and position between the two cohort groups.

[35] We also considered an alternative specification where we allowed the cohort to determine the level of efficiency units such that

$$\lambda^s_L(j, Z) = \exp(\lambda^s_{L,1} j + \lambda^s_{L,2} j^2 + I_Z Z^s_L).$$

In this simpler specification, we could consider differences in coefficient Z^s_L between finely defined cohorts since we do not run into the issue of not having observations along the support mentioned above. The coefficients for Z^s_L for cohorts differentiated by calender birth year (relative to the base cohort) all turn out to be insignificant.

Moreover, allowing for cohort-specific $\lambda_L^s(j,Z)$, $\lambda_E^s(j,Z)$ does not affect the ability of the model to match the dynamics of the relative price of experience, the age premiums and college premiums which are the focus of our analysis. The results are virtually indistinguishable from the benchmark ones, leading us to prefer the more parsimonious benchmark specification.

We can statistically assess the similarity of predictions of the models with common or cohort-specific age efficiency schedules by performing the nonparametric Kolmogorov-Smirnov test for the distributional equality between the two predicted wage distributions. Specifically, let $F_n^0(w)$ and $F_n^1(w)$ be the empirical distribution functions of the fitted log wage with sample size n from the specification of common age efficiency schedules and from that of cohort-differentiated age efficiency schedules, respectively. The Kolmogorov-Smirnov statistic for testing the equality between the two distributions is

$$K_n = \sup_w \left| F_n^0(w) - F_n^1(w) \right|,$$

where $\sup_w |\cdot|$ indicates the supremum and the empirical distribution function is defined as

$$F_n^j(w) = \frac{1}{n} \sum_{i=1}^{n} I\left(W^j \leq w\right).$$

The statistic $\sqrt{n}K_n$ converges to Kolmogorov distribution under the null of equality, which does not depend on the form of the true distribution of the log wage.

The Kolmogorov-Smirnov test statistic for the whole sample is 0.0026 with p-value of 0.639, hence we cannot reject the null hypothesis of equality between the two distributions for the whole sample. We also check the equality of the log wage distributions for the two cohorts. The test statistic for the cohort born before 1946 sample is 0.0048 with p-value of 0.439. The test statistic for the cohort born after 1946 is 0.0049 with p-value of 0.189. Thus, the equality of cohort-specific conditional distributions is not rejected either.

A6 Assessing Alternative Specifications

In this section, we compare the ability of our specification of the wage equation to fit the data relative to various alternative specifications. Two types of restrictions are of particular interest. First, our benchmark specification incorporates three potential sources of curvature in the life-cycle profiles ($g(e)$, λ_L, and λ_E). We assess whether all three components are essential for fitting the data or only a subset of them would be statistically sufficient. Second, our benchmark specification filters out the time-varying college premium, gender premium, etc. when assessing the relative price of experience, and incorporates these time varying premiums when calculating the aggregate effective supplies of experience and labor. The traditional cell-based correction for composition does not accommodate these features as we discuss in Footnote 2. Consequently, it is of some interest to assess the consequences of this restriction for fitting the wage distribution. To do so we consider alternative specifications where the coefficients on these characteristics are forced to be time-invariant.

We measure the distance between our benchmark specification and each alternative one by the difference in estimated log likelihoods between them. Vuong (1989) shows that under regularity conditions, the likelihood ratio test statistic converges to a central chi-square distribution.[36] Specifically, suppose there are two competing models to explain the variable Y conditional on Z that are represented by the conditional distribution functions $\mathbf{F}_\theta \equiv \left\{F_{Y|Z}\left(\cdot|\cdot;\theta\right);\theta \in \Theta \subset R^p\right\}$ and $\mathbf{G}_\gamma \equiv \left\{G_{Y|Z}\left(\cdot|\cdot;\gamma\right);\gamma \in \Gamma \subset R^q\right\}$, respectively, and their density functions are denoted by $f\left(y|z;\theta\right)$ and $g\left(y|z;\gamma\right)$. In our case, $y = \ln w$ and $z = (1, s, x, j, e, \chi)$. Let $\widehat{\theta}_n$ and $\widehat{\gamma}_n$ be the corresponding maximum likelihood estimators for the sample $(y_\iota, z_\iota)_{\iota=1}^n$ of size n, i.e., $\widehat{\theta}_n = \arg\max_{\theta\in\Theta} \sum_{\iota=1}^{n} \log f\left(y_\iota|z_\iota;\theta\right)$ and $\widehat{\gamma}_n = \arg\max_{\gamma\in\Gamma} \sum_{\iota=1}^{n} \log g\left(y_\iota|z_\iota;\gamma\right)$. Then, under the regularity conditions, $2LR_n\left(\widehat{\theta}_n, \widehat{\gamma}_n\right) \xrightarrow{D} \chi^2_{p-q}$, where

$$2LR_n\left(\widehat{\theta}_n, \widehat{\gamma}_n\right) = 2\sum_{\iota=1}^{n} \log\left[\frac{f\left(y_\iota|z_\iota;\widehat{\theta}_n\right)}{g\left(y_\iota|z_\iota;\widehat{\gamma}_n\right)}\right], \tag{A2}$$

and $p - q$ is the difference in the total number of parameters between the two models.

[36] See assumptions A1-A5 and information matrix equivalence condition in equation (3.8) for regularity conditions and Theorem 3.3 and Corollary 3.4 for the characterization of the asymptotic distribution of the likelihood ratio test statistic in Vuong (1989).

Table A-3: Likelihood Ratio Test Statistics.

Specification #	Description	$\mathcal{LR}$	$\chi^2_{p-q}\,(0.01)$
Spec 1	$g(e) = e$; $\lambda_L = \lambda_E = 1$; $\alpha_t = \alpha$	6943.8	181.8
Spec 2	$g(e) = e$; $\lambda_L = \lambda_E = 1$; benchmark α_t	5661.5	41.6
Spec 3	benchmark $g(e)$; $\lambda_L = \lambda_E = 1$; $\alpha_t = \alpha$	3634.5	178.4
Spec 4	$g(e) = e$; benchmark λ_L and λ_E; $\alpha_t = \alpha$	1110.8	159.0
Spec 5	benchmark $g(e)$; $\lambda_L = \lambda_E = 1$; benchmark α_t	2371.0	37.6
Spec 6	benchmark $g(e)$; benchmark λ_L and λ_E; $\alpha_t = \alpha$	1033.4	155.5
Spec 7	benchmark $g(e)$; $\lambda_L = 1$; benchmark λ_E; benchmark α_t	863.9	23.2
Spec 8	benchmark $g(e)$; benchmark λ_L; $\lambda_E = 1$; benchmark α_t	765.0	23.2
Spec 9	benchmark $g(e)$; symmetric $\lambda_L = \lambda_E$; benchmark α_t	758.9	23.2
Spec 10	$g(e) = e$; benchmark λ_L and λ_E; benchmark α_t	32.4	11.3

The alternative specifications that we consider in this section are nested by the benchmark specification. Hence the alternative hypothesis to the null hypothesis of the equivalence of the compared models is that the benchmark model is *strictly superior* to the other candidate model in fitting the wage distribution. Thus, the statistic in (A2) allows us to perform a statistical significance test for the superiority of our benchmark specification over the alternatives.

Table A-3 provides the likelihood ratio test statistic (denoted by $\mathcal{LR}$) comparing our benchmark specification with various alternatives, along with the corresponding critical values of the chi-square distributions for the 1% significance level (denoted by $\chi^2_{p-q}\,(0.01)$).

The likelihood ratio test statistics are far larger than the 1% significance critical values for all of the alternative specifications. In fact, the test statistics also exceed the 0.1% significance critical values. Our benchmark specification fits the wage distribution strictly better than the other candidate specifications at any conventional significance level. Thus, the full incorporation of age efficiency schedules for both experience and labor, the curvature of experience input, and allowing for the time-varying coefficients for control variables provides critical improvements in fitting the wage distribution.

Furthermore, the likelihood ratio test statistic can be considered as the distance of each alternative specification from the benchmark. By comparing the magnitudes of the likelihood ratio statistics across specifications, we can infer the relative importance of each ingredient of the model specification. For example, the likelihood ratio falls from 6943.8 (Spec 1) to 5661.5 (Spec 2) by allowing for time-varying coefficients on the control variables, but falls to 3634.5 (Spec 3) after relaxing the linearity of $g(e)$ function. Thus, allowing for curvature in the experience accumulation technology seems more important than allowing for the time-varying coefficients for the control variables. The most substantial improvements come from introducing the age efficiency schedules. The likelihood ratio falls from 6943.8 (Spec 1) to 1110.8 (Spec 4), after incorporating the age efficiency schedules into the model. With curvature in the experience accumulation technology and time-varying coefficients on the control variables, the likelihood ratio increases to 2371.0 (Spec 5) from 1110.8 (Spec 4) when the efficiency of labor and experience is not allowed to depend on age.

Further evidence of the importance of the full consideration of the age efficiency schedules comes from the comparison of likelihood ratios among specifications 7, 8, 9, and 10. After allowing for the age efficiency schedules and time-varying coefficients for the control variables, the likelihood ratio falls to 32.4 (Spec 10), even when restricting the experience accumulation technology to be linear. However, even with full curvature of experience accumulation technology and time-varying coefficients on the control variables, an incomplete inclusion of the age efficiency schedules makes the model fit much worse: the likelihood ratio becomes 863.9 (Spec 7 for $\lambda_L = 1$), 765.0 (Spec 8 for $\lambda_E = 1$), and 758.9 (Spec 9 for $\lambda_L = \lambda_E$).

Spec 6 shows the likelihood ratio is high (1033.4) when we do not allow for time varying premiums to college, gender etc. $\alpha_t = \alpha$. Comparing Spec 1 with Spec 2, Spec 3 with Spec 5 and Spec 4 with Spec 10, we further confirm that setting these premiums as constant under other specifications for efficiency schedules substantially raises the likelihood ratio.

This evidence implies that all three potential sources of curvature of life-cycle profiles as well as the time variation in the coefficients on the control variables are essential for fitting the wage data. We emphasize once again, however, that while this evidence guides us in specifying the model of individual earnings, it is independent of the relationship between the aggregate relative supply of experience and its relative price.

A7 Identification of the Aggregate Production Function Parameters in the Structural Model

The log wage equation in (20) includes all parameters of the model. In particular, given the measurement of aggregate inputs E_t and L_t, the variation of the relative price of experience Π_{E_t} in relation to the variation of the experience-labor ratio $\frac{E_t}{L_t}$ is the source of identification of the technology parameters μ and δ. The time-series correlation between the relative price Π_{E_t} and the relative factor endowment $\frac{E_t}{L_t}$ identifies μ (which is scale free). The average magnitude of Π_{E_t} relative to the magnitude of the $\frac{E_t}{L_t}$ identifies the scale parameter δ.

Note that the magnitudes of Π_{E_t} and $\frac{E_t}{L_t}$ depend on the normalization of some parameter of the age efficiency schedules, i.e., $\lambda_E(0, HS, x) = \lambda_L(0, HS, x) = 1$ for $x \in \{M, F\}$. Thus, the identification of δ is subject to this normalization. More precisely, it is the normalization of the *relative* efficiency of experience of the youngest workers that affects the identification of δ. That is, re-normalizing $\lambda_E(0, s, x) = \lambda_L(0, s, x) = l$ for any arbitrary constant l such that $\lambda_{E/L}(0, s, x) = 1$ leaves the estimate of δ unchanged. However, if we normalize the age efficiency schedules asymmetrically between experience and labor so that $\lambda_L(0, s, x) = a$ and $\lambda_E(0, s, x) = b$, hence $\lambda_{E/L}(0, s, x) = c = b/a \neq 1$, the coefficient function in front of experience in the log wage becomes $\delta \left(c\frac{E_t}{L_t}\right)^{\mu-1} c\lambda_{E/L}(j, s, x) = \widetilde{\delta} \left(\frac{E_t}{L_t}\right)^{\mu-1} \lambda_{E/L}(j, s, x)$, where $\widetilde{\delta} = \delta c^{\mu}$. Thus, the estimated value of δ may change. The normalization of the age efficiency schedule of *labor* affects the scale of the aggregate productivity term. Specifically, with $\lambda_L(0, s, x) = a$, the aggregate productivity term turns to $\ln aA_t$. Note, however, that estimates of μ as well as the age efficiency schedules, our key parameters, are *not* affected by this normalization.

A8 Structural Estimation, Additional Results

A8.1 Structural Estimation, Benchmark Coefficient Estimates

Table A-4: Structural Estimation, Estimates of Time-Invariant Parameters.

Parameter	Estimate	Standard error	t-statistic
$\lambda_{L,1}\,(HS, M)$	.0239	0.00335	7.13
$\lambda_{L,2}\,(HS, M)$	-.000502	0.0000846	-5.93
$\lambda_{L,0}\,(C, M)$	-.308	0.0353	-8.73
$\lambda_{L,1}\,(C, M)$	.0679	0.00366	18.52
$\lambda_{L,2}\,(C, M)$	-.00118	0.0000857	-13.82
$\lambda_{E/L,1}\,(HS, M)$	-.0773	0.00832	-9.29
$\lambda_{E/L,2}\,(HS, M)$	.00107	0.000183	5.88
$\lambda_{E/L,0}\,(C, M)$	.321	0.122	2.62
$\lambda_{E/L,1}\,(C, M)$	-0.148	0.0132	-11.19
$\lambda_{E/L,2}\,(C, M)$	0.00219	0.000266	8.25
$\lambda_{L,1}\,(HS, F)$	.000879	.00205	0.43
$\lambda_{L,2}\,(HS, F)$	.0000755	.0000467	1.62
$\lambda_{L,0}\,(C, F)$	-.0544	.0284	-1.92
$\lambda_{L,1}\,(C, F)$	.0344	.00269	12.81
$\lambda_{L,2}\,(C, F)$	-.000574	.0000622	-9.23
$\lambda_{E/L,1}\,(HS, F)$	-.0426	.00661	-6.45
$\lambda_{E/L,2}\,(HS, F)$	.0000302	.000137	.22
$\lambda_{E/L,0}\,(C, F)$	-.510	.127	-3.99
$\lambda_{E/L,1}\,(C, F)$	-.054	.0129	-4.17
$\lambda_{E/L,2}\,(C, F)$	.000158	.0000287	.55
θ_1	-0.0244	0.00462	-5.30
θ_2	.00101	0.000179	5.66
θ_3	-.0000145	0.00000231	-6.29
northeast	0.19	0.004	43.66
north central	0.046	0.004	11.49
west	0.098	0.0045	21.94
R^2	0.924		
$RMSE$	0.616		

Note - The entries represent the results of the structural estimation of time-invariant parameters of the benchmark specification. For sample restrictions and variable construction procedures, see Appendix A2. See Section 4 for details of the estimation procedure.

Table A-5: Structural Estimation, Estimates of Time-Varying Parameters.

Year	Schooling	Male	Black	Intercept
1968	.071(.005)	.25(.033)	-.16(.048)	.49(.066)
1969	.067(.004)	.24(.031)	-.16(.044)	.58(.061)
1970	.062(.004)	.24(.030)	-.14(.042)	.65(.060)
1971	.069(.004)	.22(.029)	-.10(.042)	.52(.059)
1972	.072(.004)	.21(.029)	-.12(.041)	.54(.059)
1973	.071(.004)	.25(.028)	-.087(.040)	.54(.057)
1974	.065(.004)	.23(.026)	-.100(.036)	.63(.054)
1975	.068(.004)	.21(.026)	-.086(.035)	.56(.054)
1976	.069(.004)	.16(.026)	-.051(.035)	.53(.055)
1977	.077(.004)	.19(.026)	-.037(.035)	.41(.055)
1978	.073(.004)	.19(.026)	-.052(.035)	.44(.055)
1979	.071(.004)	.20(.025)	-.086(.032)	.47(.053)
1980	.068(.004)	.19(.024)	-.044(.031)	.47(.054)
1981	.079(.004)	.17(.024)	-.092(.032)	.28(.054)
1982	.073(.004)	.16(.024)	-.093(.032)	.34(.055)
1983	.079(.004)	.12(.025)	-.075(.032)	.21(.057)
1984	.082(.004)	.11(.024)	-.082(.030)	.19(.055)
1985	.090(.004)	.16(.024)	-.078(.030)	.01(.055)
1986	.091(.004)	.13(.024)	-.12(.030)	.03(.057)
1987	.097(.004)	.11(.024)	-.15(.029)	-.03(.057)
1988	.105(.004)	.06(.024)	-.14(.029)	-.14(.059)
1989	.110(.004)	.07(.023)	-.15(.026)	-.23(.057)
1990	.114(.004)	.05(.023)	-.11(.026)	-.29(.058)
1991	.126(.004)	.04(.023)	-.071(.027)	-.47(.059)
1992	.120(.004)	.04(.023)	-.13(.026)	-.41(.059)
1993	.113(.004)	.04(.023)	-.094(.026)	-.30(.059)
1994	.125(.004)	.08(.023)	-.16(.026)	-.49(.059)
1995	.109(.004)	.08(.023)	-.12(.026)	-.24(.059)
1996	.102(.004)	.09(.023)	-.14(.026)	-.15(.059)
1997	.116(.004)	.03(.022)	-.15(.025)	-.34(.052)
1999	.110(.004)	.08(.022)	-.15(.025)	-.19(.052)
2001	.119(.004)	.07(.022)	-.087(.024)	-.24(.053)
2003	.109(.004)	.02(.022)	-.11(.024)	-.10(.052)
2005	.118(.004)	.05(.022)	-.15(.024)	-.20(.052)
2007	.120(.004)	.06(.022)	-.17(.023)	-.18(.052)

Note - The entries represent the results of the structural estimation of time-varying parameters of the benchmark specification. Standard errors are in parenthesis. For sample restrictions and variable construction procedures, see Appendix A2. See Section 4 for details of the estimation procedure.

A9 The Effect of Cohort Size on Earnings

Given the aggregate technology (3), note that from the Euler theorem

$$\begin{aligned} G_{EE} &= -\frac{L_t}{E_t} G_{EL}, \\ G_{LL} &= -\frac{E_t}{L_t} G_{EL}. \end{aligned}$$

The aggregate stocks of labor and experience in period t can be constructed as the sum of effective supplies across cohorts indexed by age j

$$\begin{aligned} L_t &= \sum_j \lambda_L(j) N_{j,t}, \\ E_t &= \sum_j \lambda_E(j) g_t(j) N_{j,t}, \end{aligned}$$

where $N_{j,t}$ denotes the cohort size and $\lambda_L(j)$, $\lambda_E(j)$, $g(j)$ denote the efficiency schedules for a representative worker in cohort j. We suppress notation for sex and schooling and omitted productive characteristics z_{jt} and hours h_{jt} for clarity. The complementarity between two cohorts j and k is given by the condition $\frac{d^2 Y_t}{dN_{j,t} dN_{k,t}} > 0$. This cross derivative is given by

$$\begin{aligned} \frac{d^2 Y_t}{dN_{j,t} dN_{k,t}} &= A_t \left[\begin{array}{c} G_{EE} \lambda_E(j) g_t(j) \lambda_E(k) g_t(k) + G_{EL} \lambda_E(j) g_t(j) \lambda_L(k) \\ + G_{LE} \lambda_L(j) \lambda_E(k) g_t(k) + G_{LL} \lambda_L(j) \lambda_L(k) \end{array} \right] \\ &= A_t G_{EL} \frac{L_t}{E_t} \left[\begin{array}{c} -\lambda_E(j) g_t(j) \lambda_E(k) g_t(k) + \lambda_E(j) g_t(j) \lambda_L(k) \frac{E_t}{L_t} \\ + \lambda_L(j) \lambda_E(k) g_t(k) \frac{E_t}{L_t} - \left(\frac{E_t}{L_t} \right)^2 \lambda_L(j) \lambda_L(k) \end{array} \right] \\ &= A_t G_{EL} \frac{L_t}{E_t} \lambda_L(j) \lambda_L(k) \left[\begin{array}{c} -\frac{\lambda_E(j) g_t(j) \lambda_E(k) g_t(k)}{\lambda_L(j) \lambda_L(k)} + \frac{\lambda_E(j) g_t(j)}{\lambda_L(j)} \frac{E_t}{L_t} \\ + \frac{\lambda_E(k) g_t(k)}{\lambda_L(k)} \frac{E_t}{L_t} - \left(\frac{E_t}{L_t} \right)^2 \end{array} \right] \\ &= A_t G_{EL} \frac{L_t}{E_t} \lambda_L(j) \lambda_L(k) \left[\frac{E_t}{L_t} - \frac{\lambda_E(j) g_t(j)}{\lambda_L(j)} \right] \left[\frac{\lambda_E(k) g_t(k)}{\lambda_L(k)} - \frac{E_t}{L_t} \right], \end{aligned}$$

using the implications of the Euler theorem above.

Since aggregate experience-labor complementarity implies $G_{EL} > 0$, cohorts are complements when the cohort specific experience-labor ratios $\frac{\lambda_E(j) g_t(j)}{\lambda_L(j)}$ and $\frac{\lambda_E(k) g_t(k)}{\lambda_L(k)}$, are respectively lower and higher than the aggregate experience-labor ratio $\frac{E_t}{L_t}$. This is because cohorts complement each other through the effect on the aggregate experience-labor ratio. When both cohort specific experience-labor ratios are either lower or higher than the aggregate ratio, they are substitutes since $\frac{d^2 Y_t}{dN_{j,t} dN_{k,t}} < 0$.

Complementarity or substitutability is stronger the larger is $\frac{\lambda_E(j)g_t(j)}{\lambda_L(j)} - \frac{E_t}{L_t}$, i.e. the more distant the cohort specific experience-labor ratios are from the aggregate ratio. For a cohort where the experience-labor ratio coincides with the aggregate ratio, i.e. $\frac{\lambda_E(j)g_t(j)}{\lambda_L(j)} = \frac{E_t}{L_t}$, the marginal product is not affected by changes in the population of other cohorts (at the margin). For the same reason, the effect of own cohort size $\frac{d^2Y_t}{d^2N_{j,t}}$ on reducing the marginal product is rising in the absolute distance of the cohort specific experience-labor ratio from the aggregate ratio, i.e. $\frac{\lambda_E(j)g_t(j)}{\lambda_L(j)} - \frac{E_t}{L_t}$.