Accounting for the Evolution of U.S. Wage Inequality*

Lutz Hendricks†

July 15, 2013

Abstract

A large literature has documented an increase of U.S. wage inequality over the past several decades. This paper asks how far one can go towards accounting for the evolution of several measures of U.S. wage inequality based on a stochastic human capital model. The model features four school levels with distinct labor types and skill prices and Ben-Porath style human capital accumulation on the job. Four exogenous changes generate rising wage inequality in the model economy: skill-biased technical change, demographic change, rising schooling, and increasing shock variances.

I find that human capital theory can account for almost the entire increase of various dispersion statistics, such as the standard deviation of log wages, the 90/50 wage ratio, and the age-specific college wage premium. The model also accounts for the changing returns to experience over time. Counterfactual experiments identify the proximate causes of rising wage inequality as well as their implications for lifetime inequality.

JEL: D31 - Personal Income, Wealth, and Their Distributions. E24 - Employment; Unemployment; Wages; Intergenerational Income Distribution; Aggregate Human Capital. I24 - Education and Inequality. J24 - Human Capital; Skills; Occupational Choice; Labor Productivity. J31 - Wage Level and Structure; Wage Differentials.

Key words: Human capital. Wage inequality.

*For helpful comments I thank seminar participants at the 2013 QSPS Workshop and PET13. The usual disclaimer applies.

†University of North Carolina, Chapel Hill; lutz@lhendricks.org

1 Introduction

The question: A large literature has documented an increase of U.S. wage inequality over that last several decades.[1] Figure 1 shows the evolution of several wage inequality measures that have been highlighted in this literature.[2] Various measures of wage inequality, including the standard deviation of log wages (panel a) and of residual log wages (panel b) have been rising since the 1970s. By contrast, the college wage premium declined during the 1970s but rose thereafter (panel c). Finally, returns to experience declined until the 1950 birth cohort and increased thereafter (panel d).

The literature has offered a variety of explanations for these observations. Following Katz and Murphy (1992), a number of authors have argued that skill-biased technical change together with changing relative labor supplies of skilled and unskilled workers accounts for the rise in the college wage premium and potentially also for at least part of the increase in wage dispersion.[3] A dissenting view attributes much of the rise in wage inequality to changes in labor market institutions, such as unionization or minimum wages.[4] Gottschalk and Moffitt (1994) argue that the variance of transitory and permanent wage shocks has increased over time, which could account for the rise in residual wage variance. Kambourov and Manovskii (2009) document declining returns to experience over time and argue that rising occupational mobility may play a role.[5]

While this literature offers important insights, it falls short of presenting a unified quantitative account of the wage distribution and its evolution. This paper offers a step in that direction. Specifically, I study how far can one go towards accounting for the evolution of U.S. wage inequality in a standard human capital model. Working with a structural model

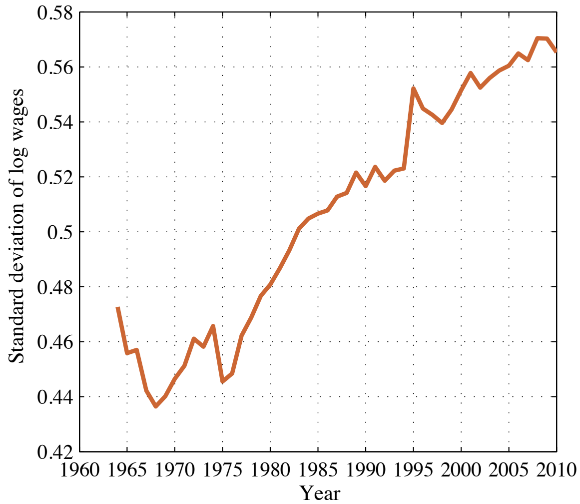
[1]For recent surveys see Katz and Autor (1999), Acemoglu (2002b), Eckstein and Nagypal (2004), and Lemieux (2007).

[2]All statistics are shown for men, aged $25-55$, and based on CPS data. Section 3 explains in detail how these measures are constructed. Similar statistics are presented, for example, in Autor, Katz, and Kearney (2008) and Heathcote, Perri, and Violante (2010).

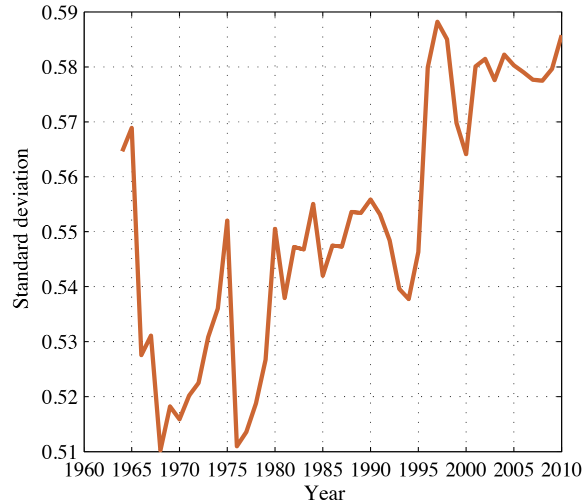
[3]For additional evidence in favor of skill-biased technical change see Berman, Bound, and Griliches (1994), Berman, Bound, and Machin (1998), and Autor, Katz, and Kearney (2008). Card and Lemieux (2001) suggest that young and old workers are imperfect substitutes in the labor market in order to account for the diverging behavior of the college wage premiums earned by both groups first documented by Katz and Murphy (1992) and Murphy and Welch (1992).

[4]See Lemieux (2007) for a summary and references.

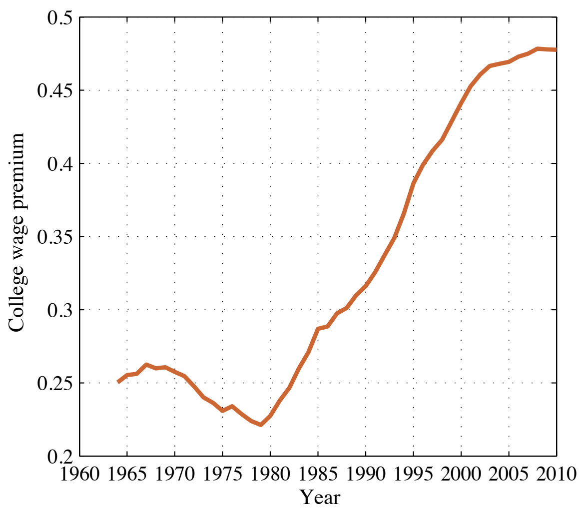
[5]Bound and Johnson (1992) and Katz and Murphy (1992) document increasing cross-sectional returns to experience.



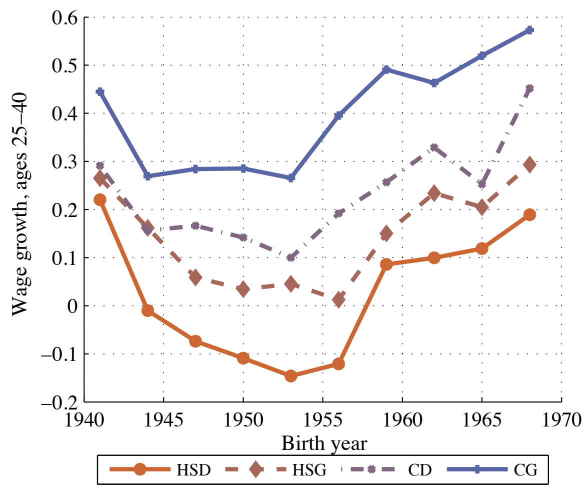
(a) Standard deviation of log wages



(b) Standard deviation of residual log wages



(c) College wage premium



(d) Returns to experience

Figure 1: Selected wage inequality statistics

Notes: Residual log wages are the residuals of a standard cross-sectional Mincerian wage regression. Returns to experience are measured by the change in mean log wages between the ages $25-40$. See Section 3 for details on the data and estimation.

allows me to trace out the indirect effects of shocks. For example, changing skill prices affect human capital accumulation, which in turn affects both between group and within group inequality. I use counterfactual experiments to quantify these indirect effects.

The approach: I develop a stochastic Ben-Porath model with school choice. The model features overlapping generations of heterogeneous workers who maximize expected lifetime earnings. As young adults, workers draw random human capital endowments and learning abilities. Next, each worker chooses from 4 discrete schooling levels (high school dropout, high school graduate, college dropout, college graduate). After schooling, the worker enters the labor market, where he accumulates additional human capital using a Ben-Porath technology. Wages are subject to persistent idiosyncratic shocks. Workers of different school groups are imperfect substitutes in production and earn different skill prices.

Four exogenous forces drive changes in the wage distribution:

1. Demographics: The age composition of the labor force changes over time.

2. Skill-biased technical change: The relative labor productivity of skilled labor grows over time. Following Katz and Murphy (1992), I assume a constant rate of skill-biased technical change.

3. School costs: The non-pecuniary costs of schooling change over time so as to account for the educational attainment of each cohort.

4. Shock variances: Following Gottschalk and Moffitt (1994), the variance of idiosyncratic productivity shocks is allowed to vary over time.

The model is parameterized to match the the following data moments:

1. the age profiles of means and standard deviations of log wages for 12 birth cohorts observed in the Current Population Survey (CPS);

2. several aggregate wage distribution statistics, estimated from CPS data;

3. the covariance structure of wages, estimated from Panel Study of Income Dynamics (PSID) data.

Results: The model successfully accounts for most of the changes in the wage distribution highlighted in the literature:

1. the increase in overall wage dispersion and its onset around 1970;
2. the differential changes in wage dispersion by schooling;
3. the rise in the college wage premium and its onset around 1970;
4. the differential increase in the college wage premium experienced by young versus old workers (Card and Lemieux, 2001);
5. the fanning out of the wage distribution (higher wage percentiles experienced faster wage growth; e.g., Autor, Katz, and Kearney 2008);
6. the U-shaped changes in the returns to experience shown in Figure 1;
7. the rising variance of residual log wages, although this is mostly a mechanical implication of rising shock variances.

The model fails to account for the fact that inequality first increases in the lower half of the wage distribution (the 50/10 ratio) and only later in the upper half (the 90/50 ratio). The literature has attributed this observation to changes in unionization and minimum wage laws, which the model abstracts from (see DiNardo, Fortin, and Lemieux 1996 and Card and DiNardo 2002).

I use counterfactual experiments to identify the proximate causes of rising wage inequality. The main findings are:

1. The rise in the college wage premium is largely due to the direct effect of changing skill prices, which are in turn caused by changes in labor supply and skill biased technical change as suggested by Katz and Murphy (1992).
2. The rise in the standard deviation of log wages is due in roughly equal parts to diverging skill prices and to rising shock variances.
3. Within school group and within school/age group inequality rises mostly due to increasing shock variances.

4. Changes in returns to experience are due to changing skill prices and to the variation in human capital investment they generate.

5. Aside from its effect on skill prices, the expansion of education makes no important contribution to changing wage inequality.

A separate contribution of the paper is to develop a stochastic Ben-Porath model that can be solved in closed form. This greatly reduces the computational burden relative to previous approaches that rely on value function approximations (e.g. Huggett, Ventura, and Yaron 2011).

1.1 Related Literature

This paper is closely related to work by Heckman, Lochner, and Taber (1998) and Guvenen and Kuruscu (2009) who also study the evolution of the U.S. wage distribution through the lens of human capital theory. Guvenen and Kuruscu's model seeks to *replace* the prevailing view that the evolution of the college wage premium is due to constant skill-biased technical change combined with changing relative labor supplies. In their environment, skilled and raw labor are perfect substitutes, so that labor supplies do not affect wages. Rising wage inequality results from an acceleration of skill-biased technical change starting in the 1970s.

Rather than replacing the prevailing view, my approach complements it and fleshes out its implications for various dimensions of wage inequality. My modeling choices therefore deviate from Guvenen and Kuruscu's:

1. I study a model with discrete school choice where workers of each education level earn a different skill price. This permits a clearer mapping between model and data because education levels are unambiguously defined.[6] It is also closer to the prevailing interpretation that the rising college wage premium reflects a change in the price of college educated labor rather than a change in the price of human capital versus raw labor. While Guvenen and Kuruscu (2009) argue that a standard Ben-Porath model cannot account for the changes observed in the data, I find that such a model can go a long way towards that objective.

[6]In Guvenen and Kuruscu's model schooling is the corner solution of the Ben-Porath human capital investment problem. A worker is college educated if he spends at least 2 years in full-time training.

2. While Guvenen and Kuruscu (2009) solve a deterministic version of their model, my model includes wage shocks. This allows me to compare the model implications to a richer set of wage dispersion statistics. Notably, my model has implications for how wage dispersion changes over the life-cycle for each model cohort.

The first paper to study the evolution of wage inequality through the lens of human capital theory is Heckman, Lochner, and Taber (1998). They interpret changes in wage inequality as the response to a one time increase in the demand for college educated labor. Both abilities and skill prices are treated as observable. Equating unobserved abilities with cognitive test score quartiles limits the importance of ability heterogeneity (see Guvenen and Kuruscu 2009). The model also abstracts from wage shocks and therefore cannot generate realistic wage variances.

Laitner (2000) offers a theoretical analysis of how the expansion of education affects the college wage premium and wage inequality. Hendricks and Schoellman (2011) study the same question quantitatively and argue that a large part of the rising college wage premium may be due to the changing composition of college educated versus high school educated workers. Huggett, Ventura, and Yaron (2011) develop a model of human capital accumulation that accounts for wage inequality over the life-cycle. I extend their model to a non-stationary economy with discrete schooling levels that command different skill prices. Guvenen, Kuruscu, and Ozkan (2011) develop a stochastic Ben-Porath model with wage shocks, but do not report the implications for wage inequality over the life-cycle. Heathcote, Storesletten, and Violante (2010) study the implications of rising wage inequality. Their model abstracts from human capital.

Lee and Wolpin (2010) study the causes of rising U.S. wage inequality in a rich structural model that abstracts from post-school human capital investment.

The paper also relates to a large literature that documents rising U.S. wage inequality and proposes explanations. This literature is too vast to even attempt a summary. The interested reader may consult the surveys cited in Footnote 1. While this literature largely relies on reduced form or accounting decompositions, this paper develops a structural model of the wage distribution. One benefit is that the indirect effects of shocks can be calculated. For example, the model reveals that variation in returns to experience may be due to the response of human capital investment to changing skill prices.

2 The Model

2.1 The Environment

I study a general equilibrium overlapping generations model. The model abstracts from physical capital. Agents can borrow and lend at an exogenous interest rate (as in a small open economy).

Demographics: In each period τ, N_τ households are born. Each lives for T periods. Denote the year in which a person of cohort c is aged t by $\tau(c,t) = c + t - 1$. Denote the cohort aged t in year τ by $c(t,\tau)$.

Endowments: Agents enter the model as young adults and draw three random endowments: a, h_1, and p_s. Learning ability a determines how efficiently the agent produces human capital in school or on the job. h_1 denotes the age 1 endowment of human capital, which I think of as produced during childhood prior to age 1. a and $\ln(h_1)$ are drawn from a joint Normal distribution

$$\begin{bmatrix} a \\ \ln(h_1) \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \bar{h}_{c,1} \end{bmatrix}, \begin{bmatrix} 1 & \rho_{a,h1} \\ \rho_{a,h1} & \sigma^2_{h1} \end{bmatrix} \right) \tag{1}$$

I normalize $\mathbb{E}(a) = 0$ and $Var(a) = 1$. The mean of $\ln(h_1)$ grows at a constant rate: $\bar{h}_{c,1} = g_{h1}(c-1)$. $\bar{h}_{1,1}$ is normalized to 0 by choosing units of h.

p_s is a vector of "psychic costs" that determines how much the individual enjoys schooling level s. p_s is drawn from a Gumbel distribution, which simplifies the school choice problem. In each period, a person works $\ell_{s,c,t}$ market hours. They can be used for work or study.

Preferences: Upon entering the labor market, individuals maximize the expected discounted present value of lifetime earnings. Equivalently, individuals maximize the present value of utility derived from consumption subject to a lifetime budget constraint with perfect credit markets. There is no need to specify the utility function. School choice is also affected by the psychic costs p_s (details below).

Human capital production: Human capital is produced in school and on the job. Agents choose from S discrete school levels. Level s lasts T_s years and results in $h_{T_s+1} = F(h_1, a, s)$ units of type s human capital at the start of work (at age $1 + T_s$). On the job, human capital is produced from human capital and study time l_t according to

$$h_{t+1} = (1 - \delta_s)h_t + A(a,s)(h_t l_t)^{\alpha_s} \tag{2}$$

where learning ability affects productivity according to $A(a,s) = e^{A_s + \theta a}$.

Shocks: Each worker's effective labor supply is given by

$$e_{i,s,c,t} = q_{i,s,c,t}\zeta_{i,s,c,t}h_{i,s,c,t}(\ell_{s,c,t} - l_{i,s,c,t}) \tag{3}$$

q denotes an autoregressive productivity shock that evolves according to

$$\ln q_{i,s,c,t+1} = \rho_s \ln q_{i,s,c,t} + \ln \xi_{i,s,c,t+1} \tag{4}$$

with $\ln \xi_{i,s,c,t} \sim N(0, \sigma^2_{\xi,s,\tau(c,t)})$ with initial condition $\ln q_{i,s,c,T_s+1} \sim N(0, \sigma_{q,s})$. $\ln \zeta_{i,s,c,t} \sim N(0, \sigma^2_{\zeta,s,\tau(c,t)})$ denotes a transitory shock. Its variance is allowed to vary over time to capture cyclical variation in wage dispersion and measurement error.

Aggregate output and skill prices: Output is produced from human capital augmented labor according to

$$Y_\tau = [G_\tau^{\rho_{CG}} + (\omega_{CG,\tau} L_{CG,\tau})^{\rho_{CG}}]^{1/\rho_{CG}} \tag{5}$$

where

$$G_\tau = \left[\sum_{s=HSD}^{CD} (\omega_{s,\tau} L_{s,\tau})^{\rho_{HS}}\right]^{1/\rho_{HS}} \tag{6}$$

Skill prices equal marginal products: $w_{s,\tau} = \partial Y_\tau / \partial L_{s,\tau}$. $L_{s,\tau}$ denotes aggregate labor supply in efficiency units ($e_{s,c,t}$) of school group s in year τ. Slightly abusing notation, this is given by

$$L_{s,\tau} = \sum_{t=T_s+1}^{T} \sum_{i=1}^{N_{c(t,\tau)}} p_{i,s,c} e_{i,s,c,t} \tag{7}$$

where $p_{i,s,c}$ is the (endogenous) probability that agent i chooses schooling level s.

2.2 Household Problem

I solve the household problem by backward induction, starting with the work phase.

Work phase: At the start of work, the agent is endowed with h_{T_s+1} units of type s human capital and with a productivity parameter $A(a, s)$. He maximizes the expected discounted present value of lifetime earnings

$$V(h_{T_s+1}, a, s, c) = \mathbb{E} \sum_{t=T_s+1}^{T} R^{-t} Y(l_t, h_t, s, c, t) \tag{8}$$

subject to the law of motion for h (2) and the time constraint $0 \leq l_t \leq \bar{l}\ell_{s,c,t}$. R denotes the exogenous gross interest rate. $\bar{l}$ denotes the maximum fraction of time that may be devoted to human capital accumulation. Period earnings are given by

$$Y(l_t, h_t, s, c, t) = w_{s,\tau(c,t)} e_{s,c,t} \tag{9}$$

Wages are given by $y(l_t, h_t, s, c, t) = Y(l_t, h_t, s, c, t)/\ell_{s,c,t}$. Human capital investment is chosen before the current transitory shock, ζ_t, is realized.

Backward induction leads to the following condition characterizing optimal human capital investment:

$$(h_{i,s,c,t} l_{i,s,c,t})^{1-\alpha_s} = \frac{\alpha A(a, s)}{(1 - \delta_s)} \sum_{j=1}^{T-t} X_{s,c,t,j} \frac{\mathbb{E}(q_{i,s,c,t+j}|t)}{q_{i,s,c,t}} \tag{10}$$

where

$$X_{s,c,t,j} = \left(\frac{1 - \delta_s}{R}\right)^j \frac{w_{s,\tau(c,t+j)}}{w_{s,\tau(c,t)}} \ell_{s,c,t+j} \tag{11}$$

is common to all workers. Since q is governed by (4), the expected value of future shocks is given by

$$\mathbb{E}(q_{i,s,c,t+j}|t) = (q_{i,s,c,t})^{\rho^j} \times \prod_{i=1}^{j} \mathbb{E}\{(\xi_{i,s,c,\tau(c,t+i)})^{\rho^{j-i}}\} \tag{12}$$

where $\mathbb{E}\{(\xi_{i,s,c,\tau(c,t+i)})^{\rho^{j-i}}\} = \exp\left(\frac{[\sigma_{\xi,s,\tau(c,t+i)}\rho^{j-i}]^2}{2}\right)$.

The model can now be solved recursively as follows. Given worker endowments, (10) solves for l_{T_s+1} at the start of the worker's career. In case of a corner solution, l_{T_s+1} is set to $\bar{l}\ell_{s,c,T_s+1}$. The human capital technology yields h_{T_s+2}. Each worker then draws q_{T_s+2} and updates his beliefs about future shocks. Iterating forward until the age of retirement unfolds the worker's life-cycle wage profile.

School phase: At the start of life, the agent chooses schooling to maximize

$$W_s(p_s, h_1, a, s, c) = \underbrace{\ln V(F[h_1, a, s], a, s, c)}_{\text{lifetime earnings}} + \mu_{s,c} + \underbrace{\pi p_s + \pi_a(T_s - T_1)a}_{\text{"psychic cost"}} \tag{13}$$

In addition to expected lifetime earnings V the agent enjoys an idiosyncratic "psychic" utility, $\pi p_s + \pi_a(T_s - T_1)a$, and a common, school specific utility $\mu_{s,c}$. The common utility $\mu_{s,c}$ allows the model to match the fraction of persons choosing each school level in each cohort. The psychic utility plays its usual role as a stand-in friction that generates imperfect school sorting by ability and h_1 (see Heckman, Lochner, and Todd 2006 for a discussion). $\pi > 0$ is a scale factor that defines the units of the psychic cost. The term $\pi_a(T_s - T_1)a$ permits the model to generate a positive correlation between schooling and ability.[7] Since p_s is drawn from a Gumbel distribution, the probability of choosing school level s is given by $X_s / \sum_j X_j$ where

$$X_s = \exp\left(\frac{V(F[h_1, a, s], a, s, c) + \mu_{s,c} + \pi_a(T_s - T_1)a}{\pi}\right) \tag{14}$$

(e.g., Greene 2012, ch. 18.2).

2.3 Cognitive Test Scores

To help with the identification of model parameters that relate to the dispersion of abilities (θ) and to the strength of the association between ability and schooling (π, π_a) it is helpful to employ an observable proxy for ability. Following Hendricks and Schoellman (2011), I treat cognitive test scores (IQs, for short) as noisy measures of a:

$$IQ = a + \sigma_{IQ}\varepsilon_{IQ} \tag{15}$$

where $\varepsilon_{IQ} \sim N(0, 1)$ reflects measurement error. Note that IQ does not have units, so that its standard deviation is meaningless. Since ability is drawn from a standard Normal distribution, $1/\sigma_{IQ}$ is the signal-to-noise ratio.

[7] Alternatively, one could allow psychic costs and abilities to be correlated. However, this would be less tractable.

2.4 Discussion of Model Assumptions

The model is computationally very efficient. Since agents maximize expected lifetime earnings, the job-training part of the model can be solved in closed form. Only the expectation in (8) needs to be evaluated numerically. This is done by regressing realized lifetime earnings on individual endowments for a large number of simulated model households.[8] The presence of Gumbel distributed preference shocks allows the school choice problem to be solved in closed form, even if the role of the preference shocks (π) is small. It also guarantees continuity of the simulated model data with respect to the model parameters.

The optimality condition (10) reveals that workers invest more in human capital on the job when skill prices grow faster. The level of skill prices is irrelevant. Worker ability affects the level of investment, but not its sensitivity to skill prices. These features contrast with Guvenen and Kuruscu (2009)'s model where on-the-job investment depends on relative skill prices and where the sensitivity of investment to skill prices increases with ability. This interaction drives Guvenen and Kuruscu's main results. Skill-biased technical change raises the relative price of human capital, which increases investment among highly able young workers. This leads initially to a compression of the wage distribution and to a decline in the skill premium, especially among the young (Caselli, 2009). Even though both models account for a similar set of observations, the underlying mechanism is very different.

2.5 Computational Issues

Aggregation: Aggregate wage statistics for each year are computed from simulated wage histories for $1,000$ model agents in each school group. When computing statistics that aggregate ages, such as the mean log wage of college graduates in year τ, individual observations are weighted to remove compositional effects. Specifically, each person's weight is given by his probability of choosing school level s times a factor that equates the fraction of persons in each age-school cell to the corresponding cross-year average fraction in the CPS sample.

The model does not have a direct counterpart for each birth cohort in the data. The reason is that the oldest model cohort is born in 1935, which ensures that most of its wage history is observed, but it also implies that there are no old model agents in the first data years.

[8] The accuracy of this approach is verified by simulating a large number of households with identical endowments and comparing their mean lifetime earnings with the predicted value from the regression.

In computing aggregates, I assume that cohorts without model counterparts are identical to the closest model cohort.

For the same reason, I compute $L_{s,\tau}$ as the ratio of total earnings in cell (s, τ) in the CPS to $w_{s,\tau}$ as implied by the model. This ensures that the model exactly matches observed wage bills. Computing aggregate labor inputs from simulated individual work histories yields similar results.

Skill prices: The model can be solved in two ways.

1. Relative skill weights $(\omega_{s,\tau}/\omega_{HSG,\tau})$ can be restricted grow at constant rates. The model then quantifies the implications of constant skill-biased technical change and demographic change, as proposed by Katz and Murphy (1992). The baseline model follows this approach.

2. Skill weights can be left unrestricted. Skill prices are then chosen to best fit all calibration targets, including mean log experience wage profiles for all cohorts. This imposes strong discipline on the model. In particular, it does not permit the model to generate wage inequality by generating counterfactual mean wage profiles or by choosing model parameters that are not consistent with observed experience wage profiles.

Both approaches turn out to yield very similar results, supporting Katz and Murphy's (1992) claim that a combination of constant skill-biased technical change and demographic change can account for much of the rise in U.S. wage inequality.

3 Data

The data for wages, hours worked, and schooling are taken from the March Current Population Survey (CPS) for the years $1964-2011$ (King, Ruggles, Alexander, Flood, Genadek, Schroeder, Trampe, and Vick, 2010).[9] The sample contains men born between 1935 and 1968. To increase sample sizes, I divide the population into 12 equally spaced birth cohorts.

[9] Relative to the Panel Study of Income Dynamics, the main advantage of CPS data is the much larger sample size. The drawback is that individuals cannot be followed over time.

Each cohort pools 3 adjacent birth years. For example, the cohort labeled 1950 contains persons born between 1949 and 1951.

For each cohort and school group, I construct age profiles of mean hours worked, mean log wages, and the standard deviation of log wages. The age hours profiles measure time endowments. Specifically, I set $\ell_{s,c,t}$ to mean annual hours in each (s, c, t) cell divided by 2,000, which yields time endowments for middle aged men that are close to 1. The age-wage profiles form the calibration targets, as described in Section 4.

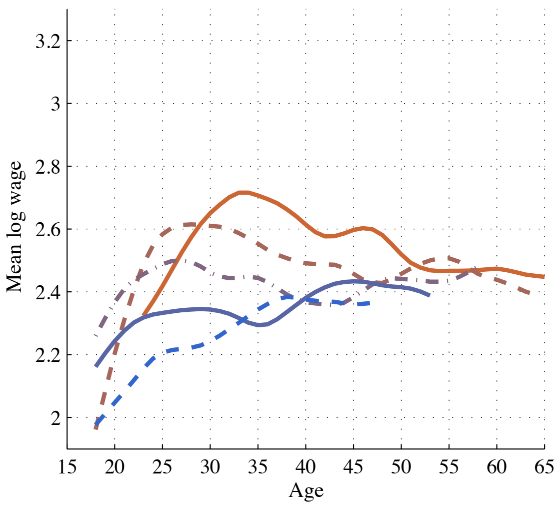
The measurement of wages closely follows Bowlus and Robinson (2012) and is described in Appendix (A). Each cohort's log wage profile is smoothed using an HP filter with parameter 5. Figure 2 displays selected cohort age-wage profiles, deflated by the consumer price index with base year 2000.[10] It is apparent that some of the profiles are not consistent with human capital theory and constant skill prices (Bowlus and Robinson 2012 make a similar observation). Some of the profiles are essentially flat past age 25. Many are not concave. Note also that the longitudinal wage profiles look very different from cross-sectional profiles that are sometimes used in their stead (see Thornton, Rodgers, and Brookshire 1997). They also look very different from the age-wage profiles that sometimes estimated using panel data and imposing that a fixed age profile, combined with either year effects or cohort effects, characterizes all cohorts (e.g., Figure 3 in Huggett, Ventura, and Yaron 2006).

3.1 Cognitive Test Scores

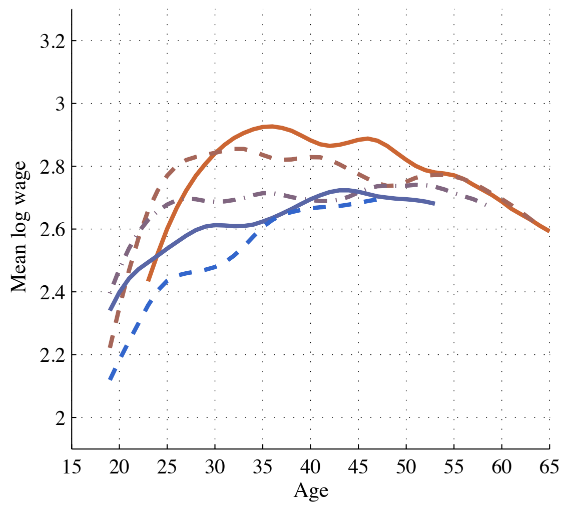
The IQ related data points are constructed from NLSY79 data (see Appendix 6 for details). Armed Forces Qualification Test (AFQT) scores are the empirical counterparts of IQ. The percentile scores are age-adjusted using a regression of AFQT percentiles on a quadratic in the age at which the test was taken. The scores are then transformed to have a standard Normal distribution.

To characterize the relationship between AFQT and wages, I regress log wages on AFQT scores, AFQT scores interacted with a cubic polynomial in experience, as well as controls for school group, years of schooling, experience, marital status, race, and region of

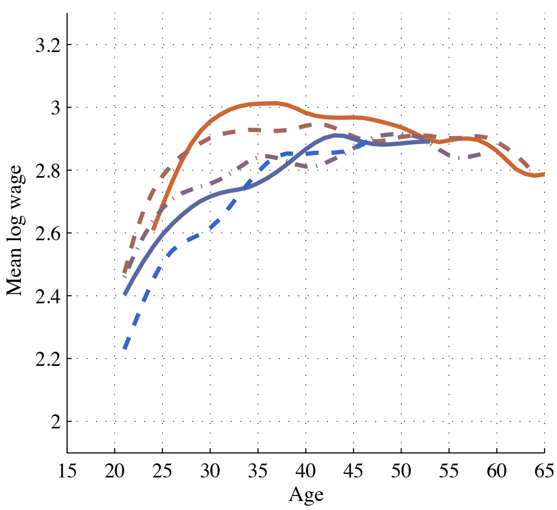
[10] Bowlus and Robinson (2012) report log median wage profiles for a different set of cohorts. My data imply similar profiles for their cohorts, even though I use mean log wages instead. I am grateful to Chris Robinson for providing me with his data to perform this comparison.



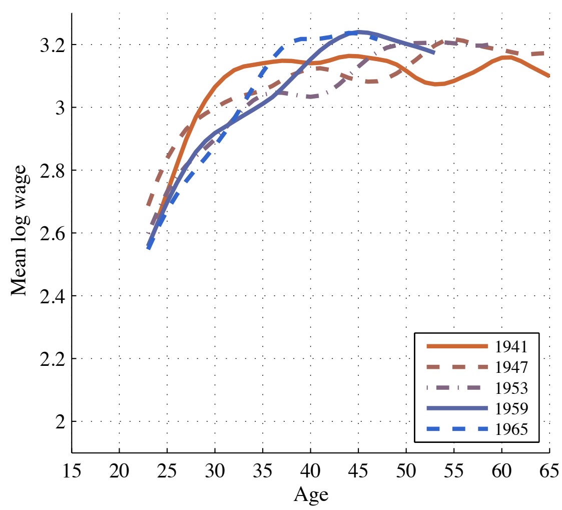
(a) High school dropouts



(b) High school graduates



(c) College dropouts



(d) College graduates

Figure 2: Age profiles of mean log hourly wages for selected cohorts

Notes: In year 2000 prices. CPS data.

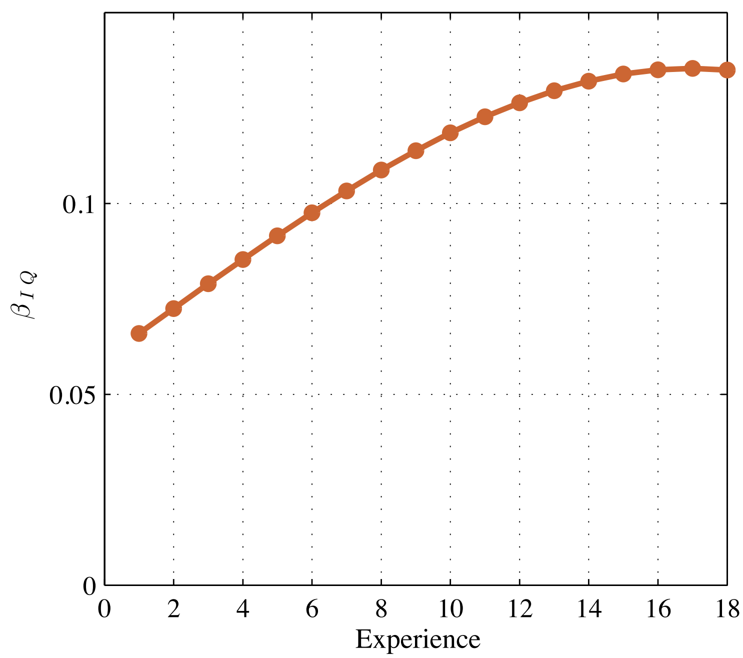


Figure 3: β_{IQ} and experience

residence. Figure 3 shows how the regression coefficient of AFQT scores varies with experience. Experience is truncated at 18 years, which is 1.5 times average experience in the data. Consistent with Altonji and Pierret (2001), β_{IQ} increases with experience. To estimate β_{IQ} in the model, I regress simulated log wages for each experience level on standard Normal IQ scores.

To characterize the strength of school sorting by ability, I draw on the average cognitive test scores for college and high school students of various cohorts. The data are taken from Taubman and Wales (1972) and described in more detail in Hendricks and Schoellman (2011).

4 Setting Model Parameters

Table 1 summarizes the parameter values that are fixed exogenously. Based on McGrattan and Prescott (2000) the gross interest rate is set to $R = 1.04$. Agents are allowed to spend at most $\bar{l} = 0.9$ of their time endowment on job training. In the data, I assume that agents start working at the ages of 18 (HSD), 19 (HSG), 21 (CD), and 23 (CG) and retire at age

Table 1: Fixed parameters

Parameter	Description	Value
T	Lifespan	65
τ	Birth cohorts	1935, 1938, 1941, ..., 1962, 1965, 1968
T_s	School duration	2, 3, 5, 7
$\ell_{t,s,\tau}$	Market hours	CPS data
n/a	Nodes of skill price spline	1950, 1957, 1964, ..., 2010, 2021, 2032
R	Gross interest rate	1.04

65, which yields the values of T_s.

For the baseline model, I also impose the following restrictions:

1. $g_{h1} = 0$: The distribution of human capital endowments is constant over time.

2. $\sigma_{\xi,s,\tau}$: The standard deviation of the persistent shock is assumed to follow a linear time trend over the data period. For years outside of the data period, $\sigma_{\xi,s,\tau}$ is set to the nearest value in the data period.

3. $\omega_{s,\tau}$: All skill weights grow at the same rate (constant skill-biased technical change).

4.1 Calibrated Parameters

The schooling technology $F(h_1, a, s)$ is the same as the job-training technology with $l_t = 1$. The following parameters are calibrated jointly:

- endowment parameters: θ, g_{h1}, π, π_a, ρ_{a,h_1}, σ_{IQ};

- human capital technologies: α_s, δ_s, A_s;

- final output technology: ρ_{HS}, ρ_{CG};

- school costs $\mu_{s,c}$, where $\mu_{1,c}$ may be normalized to 0;

- shocks: $\sigma_{q1,s}, \rho_s, \sigma_{\xi,s,\tau}, \sigma_{\chi,s,\tau}$;

- skill bias parameters: $\omega_{s,\tau}$

These parameters are jointly calibrated for all school groups using a simulated method of moments. I search over the parameter space. For each parameter guess, I solve the model and simulate $1,000$ individuals. The values of $\mu_{s,c}$ are chosen so that the model exactly matches the fraction of persons choosing each school level in each cohort. The correlation of repeated IQ tests imply a lower bound of $\sigma_{IQ} \geq 0.5$ (see Hendricks and Schoellman 2011 for details).

Iterating directly over the means and growth rates of skill bias parameters would be computationally costly. It would also preclude any short-term deviations from constant relative skill bias growth. I therefore search over skill prices ($w_{s,\tau}$) instead, which are computed from a cubic spline with 9 nodes. After computing aggregate labor supplies from the household problem, I derive the implied skill weights. Penalties for deviations from constant growth paths of $\omega_{s,\tau}/\omega_{HSG,\tau}$ force the model to approximate constant skill-biased technical change.

The household solution depends on skill prices outside of the period observed in CPS data. I construct these skill prices as follows. I assume that households who are born before the first model cohort or after the last model cohort are identical to the nearest model cohort. This restriction allows me to compute labor supplies during all periods where model households work. I allow the calibration algorithm to freely choose out of sample skill prices, but penalize deviations from the average growth rate of skill weights during the sample period.

Note that the solution to the household problem does not depend on the variance of transitory shocks. To estimate this variance, I therefore proceed as follows. I solve the model for given values of all parameters, except $\sigma^2_{\zeta,s,\tau}$. I regress the model generated standard deviation of log wages in each (age, year) cell ($\sigma_{s,c,t}$) on their data counterparts ($\sigma^d_{s,c,t}$) and on time dummies, which capture $\sigma^2_{\zeta,s,\tau}$. This is done separately for each school group. I then draw realizations of ζ_t and add them to each household's model wages and recompute $\sigma_{s,c,t}$ including the transitory shocks.

Once these restrictions are imposed, there are 68 calibrated parameters of which 36 characterize the evolution of skill prices. The algorithm minimizes the sum of squared deviations between moments computed from simulated wage histories and moments taken from the data. The moments characterize:

- age profiles of mean log wages;
- age profiles of the standard deviation of log wages;

- aggregate wage dispersion measures for each year: the college wage premium; the variance of log wages; the 90/50, 50/10 and 90/10 log wage gaps; the standard deviation of residual wages;
- the covariance matrix of log wages, estimated from PSID data (see Appendix (B));
- the age profile of β_{IQ} for the 1960 cohort, estimated from NLSY79 data;
- mean cognitive test scores for selected cohorts of high school and college students, taken from Taubman and Wales (1972).

Table 2 shows the values of the calibrated parameters. The curvature parameters of the human capital production functions lie near 0.45 which is far lower than recent estimates. Heckman, Lochner, and Taber (1998) and studied following their approach (Taber 2002; Kuruscu 2006) estimate α_s above 0.9. The reasons for the discrepancy are discussed in Hendricks (2012).

The model implies that skill types are highly substitutable. In the data, the relative supply of college educated labor rises dramatically over time, in part due to the expansion of education and in part due to the baby boom (Katz and Murphy, 1992). At the same time, the college wage premium increases. Reconciling these facts requires either very rapid skill-biased technical change or a high elasticity of substitution between college graduates and other workers.

Figure 4 shows the calibrated skill prices and compares them with observed mean log wages. Skill prices grow more slowly than observed wages, indicating that human capital grows over time. Skill prices are also far smoother than observed wages. The model avoids large predictable changes in skill prices because they would lead to fluctuations in training time and cohort age-wage profiles.

4.2 Model Fit

Figure 5 shows model generated and observed mean log age-wage profiles for high school graduates of selected cohorts. Figure 6 shows the same information for college graduates. In both cases the model replicates how the curvatures and slopes of the profiles change over time. The main discrepancies occur at young ages where the model sometimes fails to match the observed low wages earned by workers with little experience.

Parameter	Description	Value
On-the-job training		
A_s	Productivity	0.14, 0.12, 0.15, 0.23
α_s	Curvature	0.43, 0.38, 0.42, 0.48
δ_s	Depreciation rate	0.053, 0.041, 0.051, 0.089
Endowments		
σ_{h1}	Dispersion of h_1	0.297
θ	Ability scale factor	0.115
π_1	Psychic cost scale factor	0.296
γ_{ap}	Ability weight in psychic cost	0.123
γ_{ah}	Governs correlation of $\ln h_1$ and a	0.362
σ_{IQ}	Noise in IQ	0.587
Shocks		
$\sigma(q_1)$	Std dev of first shock	0.00, 0.00, 0.00, 0.00
ρ_s	Shock persistence	0.98, 0.97, 0.97, 0.97
$\sigma(\zeta)$, 1964	Std deviation of shocks	0.12, 0.11, 0.11, 0.11
$\sigma(\zeta)$, 2010	Std deviation of shocks	0.11, 0.16, 0.15, 0.14
Other		
Δw_s	Skill price growth rate, 1964-2010 [pct]	-1.00, -0.76, -0.69, -0.09
$(1+\rho_{HS})^{-1}$, $(1+\rho_{CG})^{-1}$	Substitution elasticities	8.30, 6.05

Table 2: Calibrated parameters

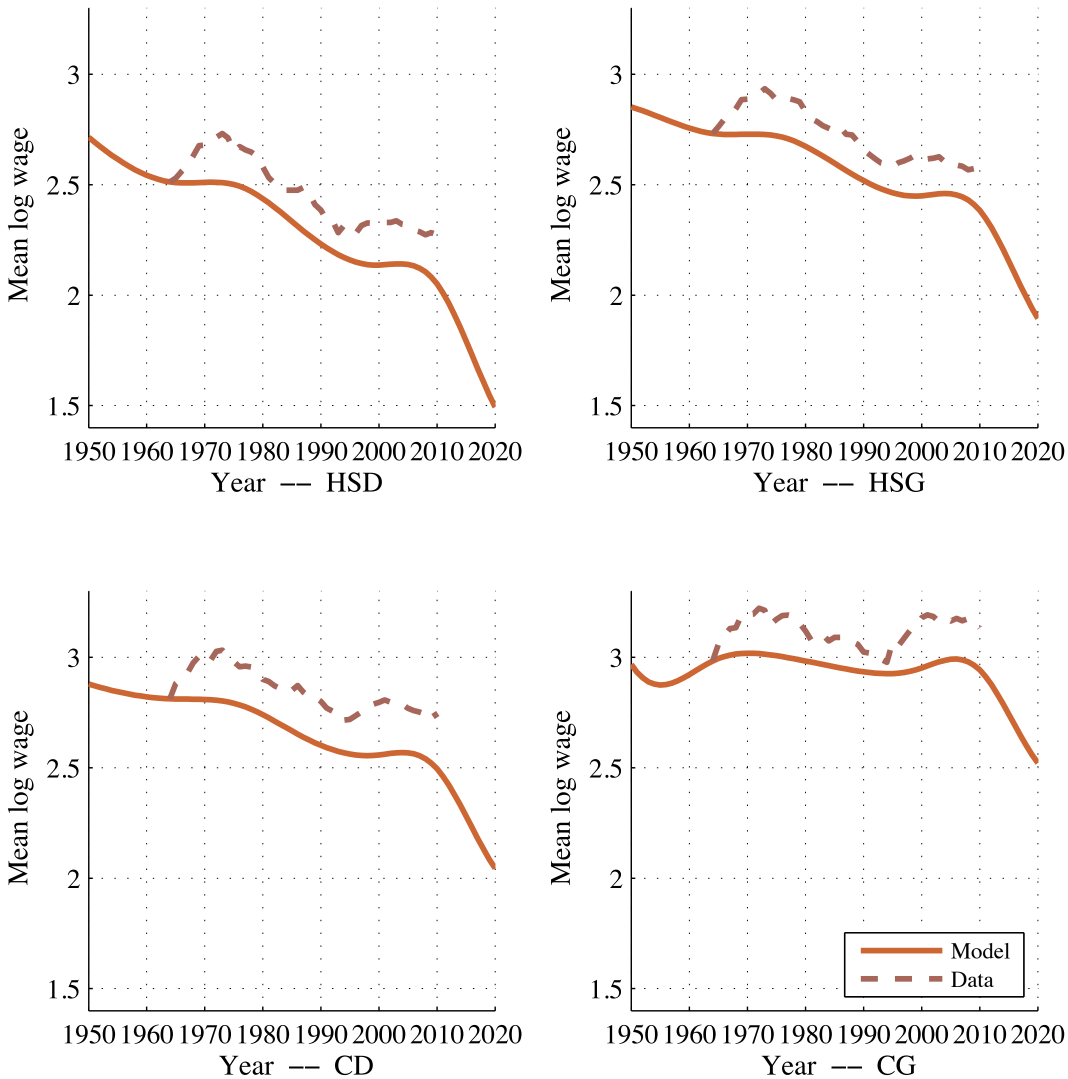


Figure 4: Skill prices and mean log wages

Notes: The figure shows calibrated skill prices and observed mean log wages. Skill prices are normalized to equal mean log wages in 1964.

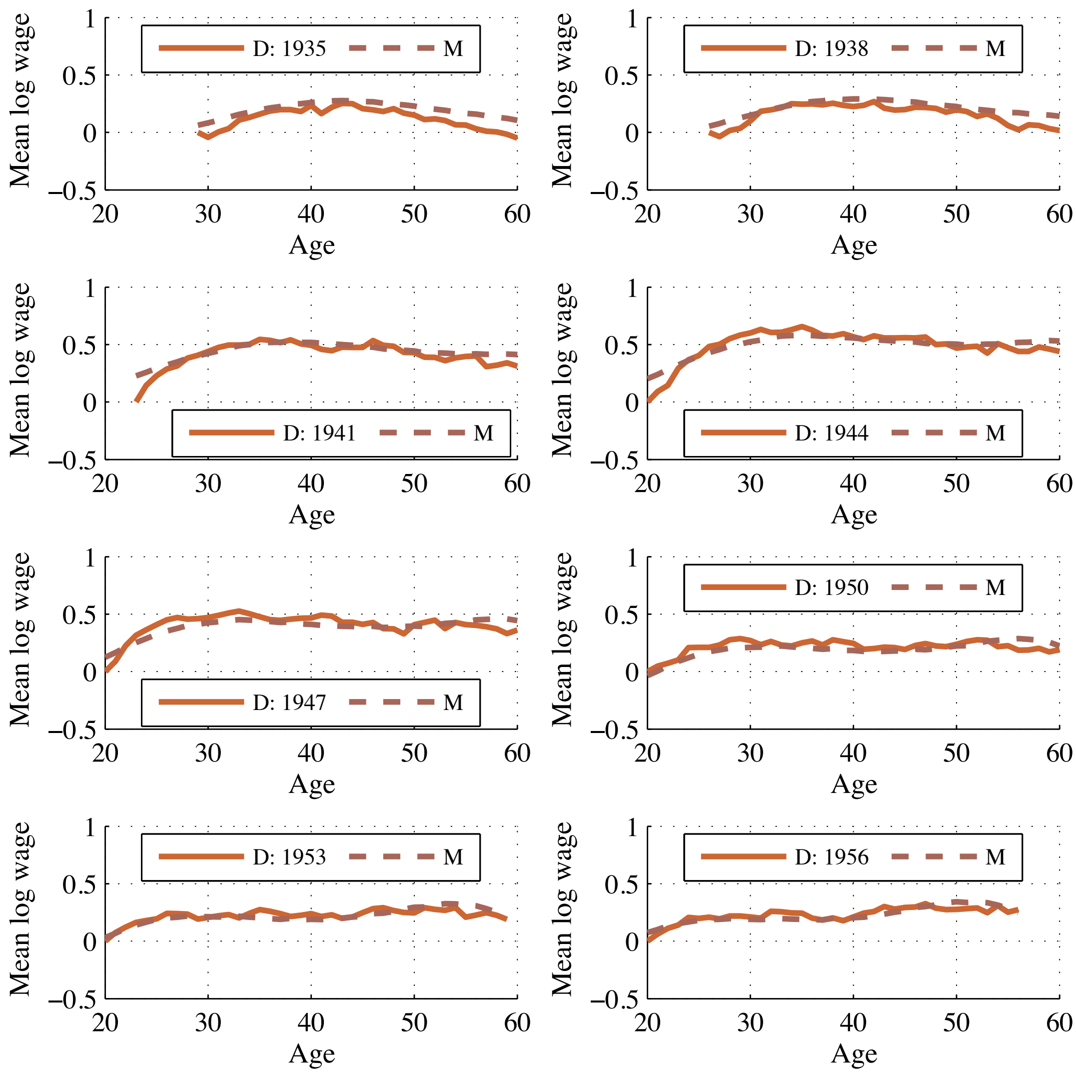


Figure 5: Mean log wages: high school graduates

Notes: The figure shows mean log wage profiles for selected cohorts. "D" followed by the cohort's birth year indicates the data profile. "M" indicates a model profile.

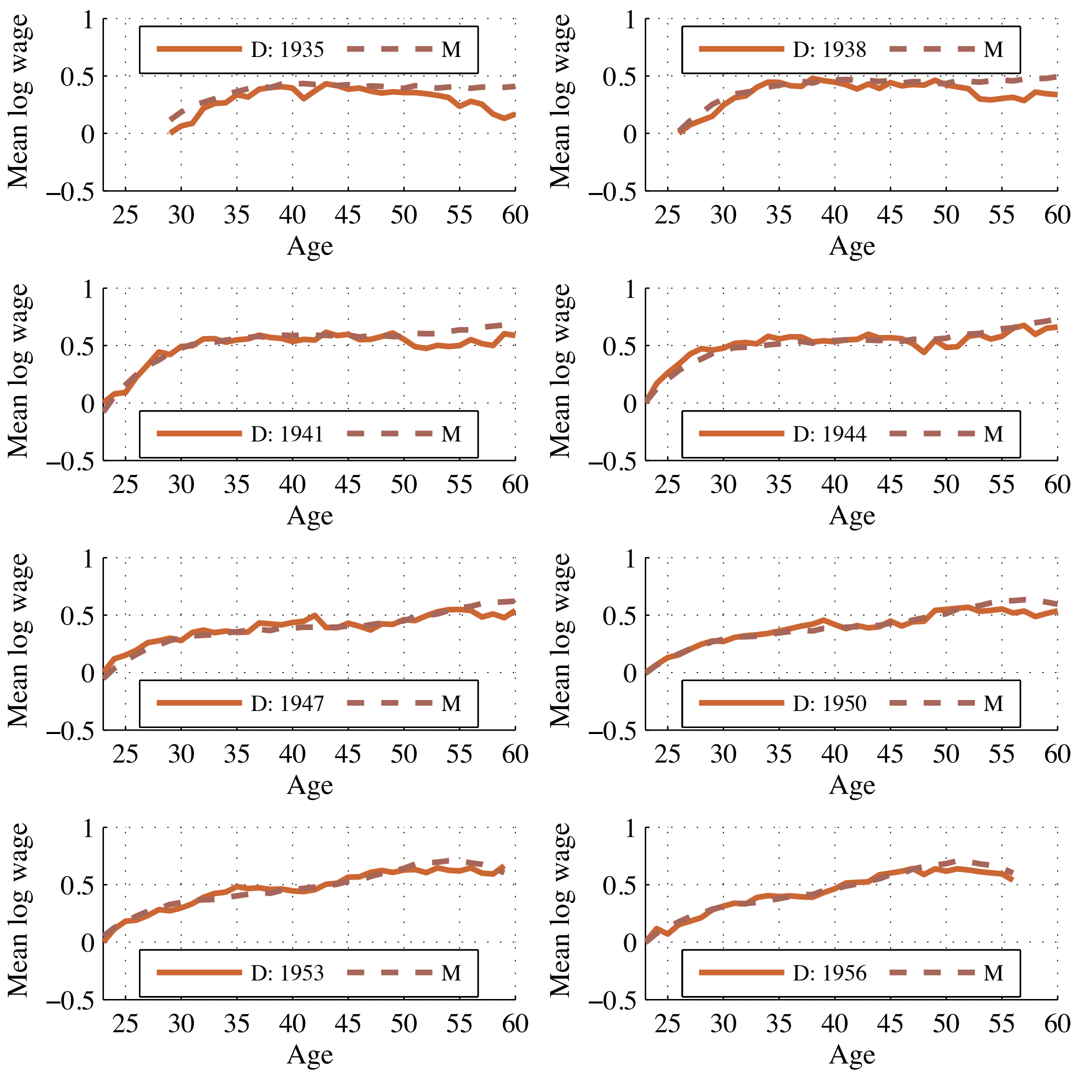


Figure 6: Mean log wages: college graduates

Notes: See Figure 5.

Figures 7 and 8 compare the age profiles of the standard deviation of log wages of model and data cohorts. For college graduates the fit is quite close. For high school graduates, the model overstates wage dispersion among the early cohorts.

Cognitive test scores: Figure 9 shows that the model closely replicates the relationship between wages and test scores as measured by β_{IQ} as well as the average test scores of high school and college students of selected cohorts.

Wage shocks: Table 3 assesses how well the model replicates the covariance structure of wages estimated from PSID data. Rather than showing the entire covariance matrix for each school group, I estimate an autoregressive process of the form

$$\ln y_{i,t} = \alpha_i + g(t) + \nu_{i,t} \tag{16}$$

$$\nu_{i,t} = \rho_s \nu_{i,t-1} + \zeta_{i,t} \tag{17}$$

$$\zeta \sim N(0, \sigma_{\zeta,s}) \tag{18}$$

by minimizing the sum of squared deviations between the covariance matrix of wages implied by this process and the one generated by either the model of PSID data. Following Guvenen (2007), workers are included in the regression if they are observed between 1967 and 1992 at ages between 20 and 60. To increase sample sizes, the empirical estimates pool high school dropouts with high school graduates and college dropouts with college graduates.

The model parameters are roughly in line with empirical estimates found in the literature. Notably Guvenen (2007) reports persistence parameters $\rho_{HSG} = 0.972$ (s.e. 0.023) and $\rho_{CG} = 0.979$ (s.e. 0.055) with standard deviations of the persistent shocks of $\sigma_{\zeta,HSG} = 0.105$ (s.e. 0.084) and $\sigma_{\zeta,CG} = 0.100$ (s.e. 0.114). The main discrepancy relative to the data is the high persistence parameter for college graduates.[11]

4.3 Skill-biased Technical Change

A leading explanation for the rising college wage premium relies on a combination of skill-biased technical change and changing labor supplies (Katz and Murphy, 1992). The model presented here is entirely consistent with this explanation.

[11] However, some empirical estimates of shock persistence imply near permanent shocks. One example is Storesletten, Telmer, and Yaron (2004).

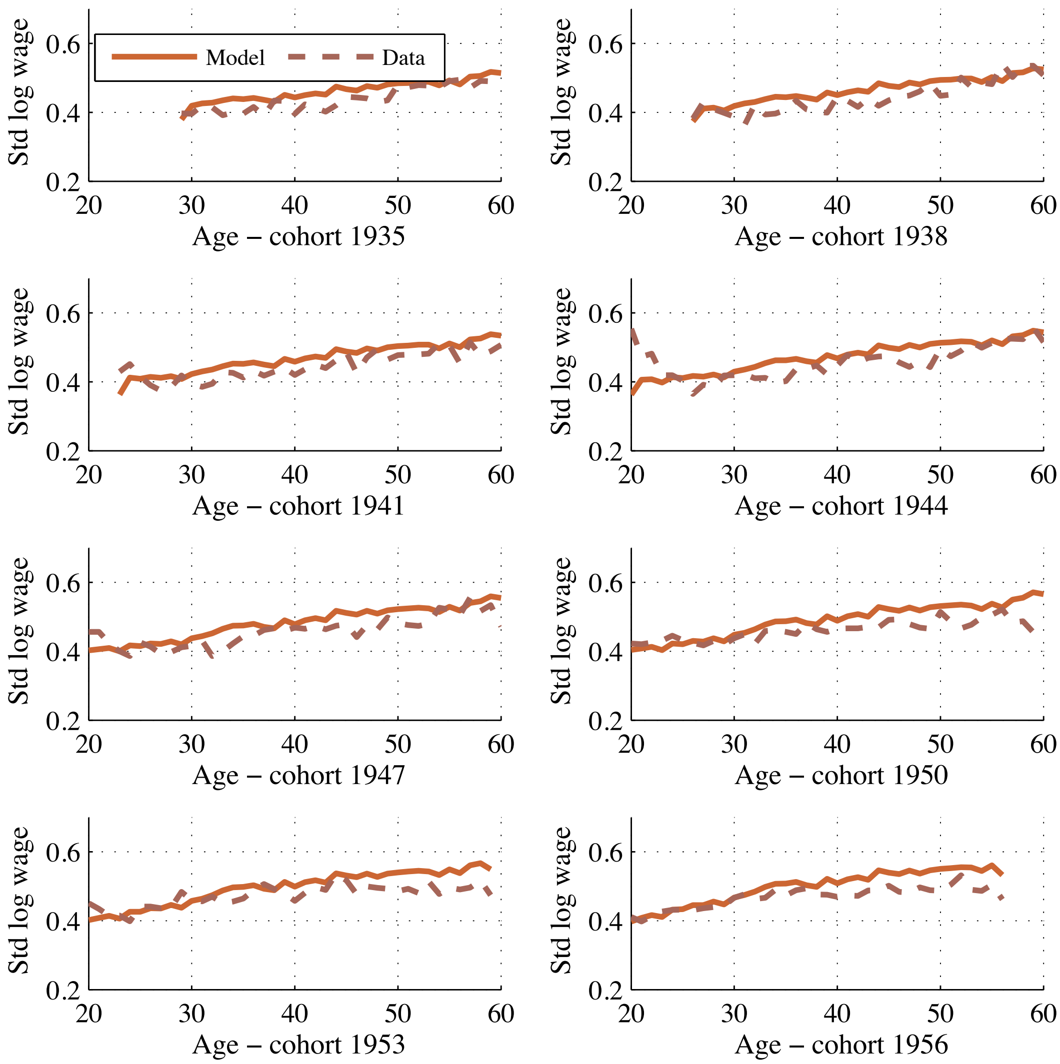


Figure 7: Model fit: high school graduates

Notes: See Figure 5.

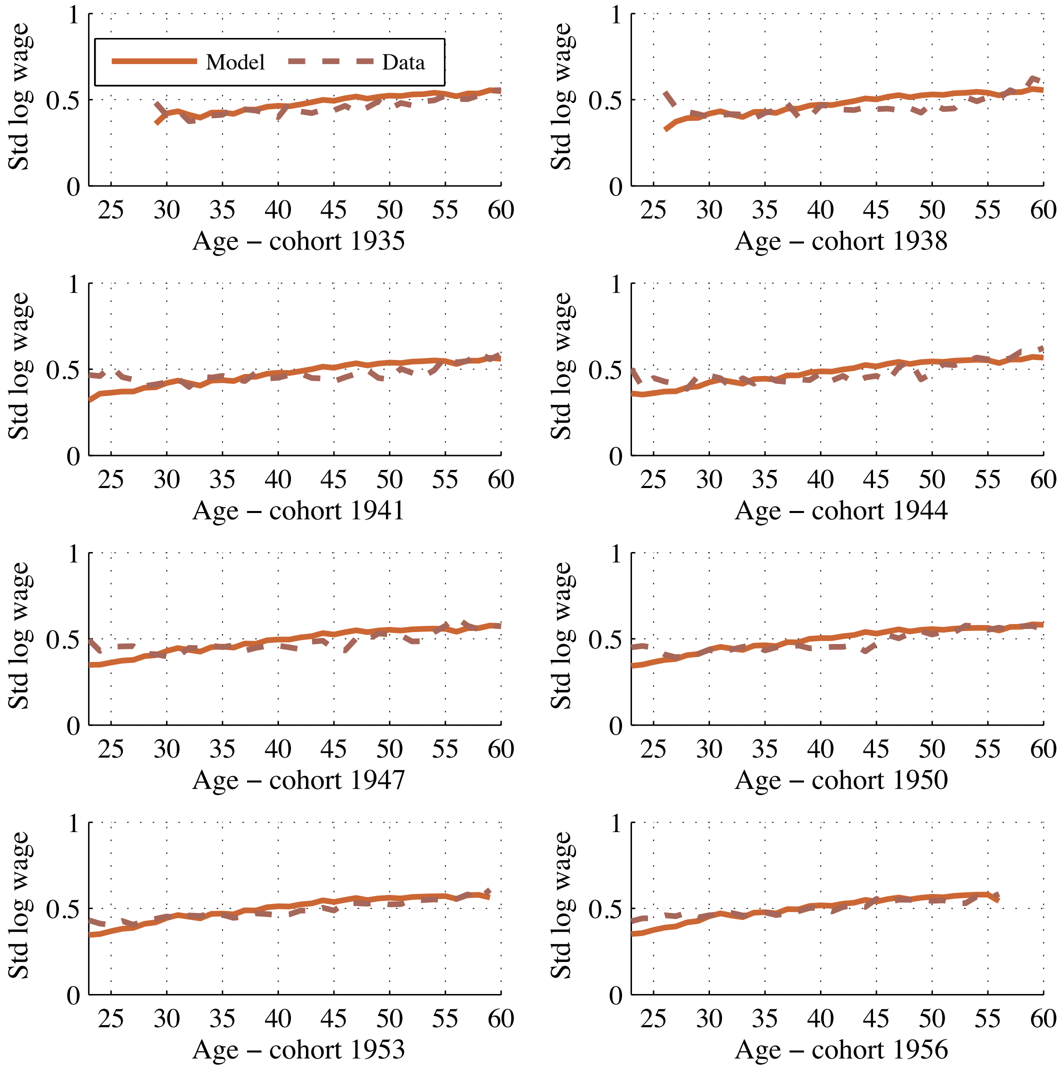
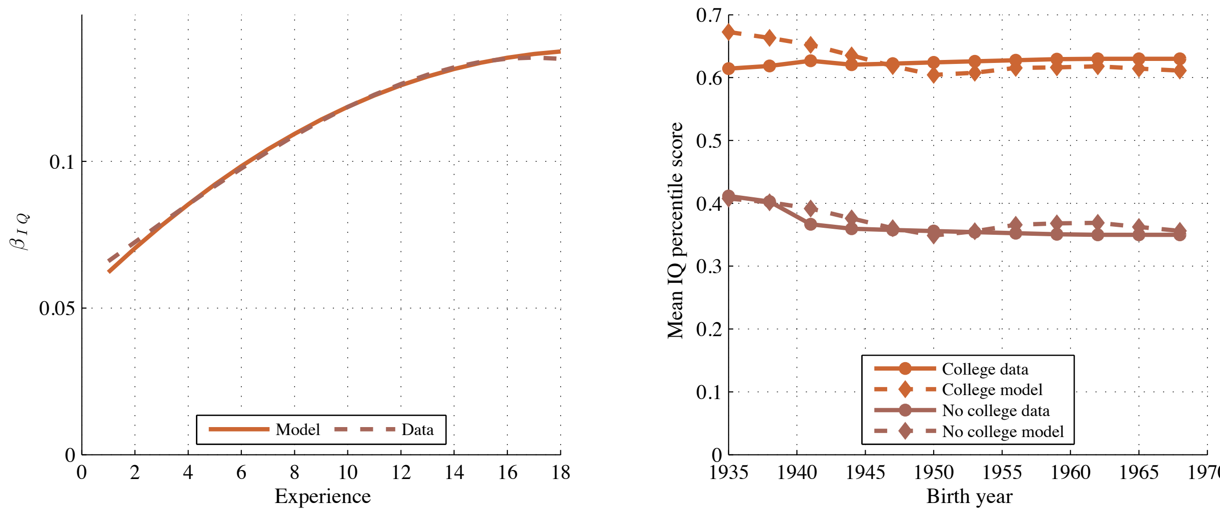


Figure 8: Model fit: college graduates

Notes: See Figure 5.



(a) β_{IQ} and experience

(b) Average test scores by cohort

Figure 9: Model fit: cognitive test scores

School / Parameter	$var(\alpha)$	ρ	$var(\eta)$	$var(\epsilon)$
HSD	0.049 (0.006)	0.955 (0.003)	0.020 (0.001)	0.029 (0.001)
Data	0.000	0.972	0.011	0.000
Model parameters		0.982	0.014	
HSG	0.046 (0.005)	0.934 (0.004)	0.024 (0.001)	0.029 (0.001)
Data	0.000	0.972	0.011	0.000
Model parameters		0.967	0.019	
CD	0.041 (0.004)	0.950 (0.003)	0.021 (0.001)	0.032 (0.001)
Data	0.000	0.979	0.010	0.000
Model parameters		0.973	0.017	
CG	0.034 (0.004)	0.954 (0.003)	0.021 (0.001)	0.035 (0.001)
Data	0.000	0.979	0.010	0.000
Model parameters		0.974	0.017	

Table 3: Wage shocks

Notes: The table shows the parameters of autoregressive wage processes that are estimated on model and PSID workers. For each parameter, the model shows the PSID estimate followed by the model estimate.

Figure (10) shows the evolution of the skill weights relative to high school graduates, $\omega_{s,\tau}/\omega_{HSG,\tau}$ (panel b). Consistent with Katz and Murphy (1992), relative skill weights grow at roughly constant rates over the data period. Panel (a) shows the skill weight levels, $\omega_{s,\tau}$, and compares them with a calculation that abstracts from human capital and measures labor inputs as total hours worked in each (s, τ) cell. Both models imply very similar results

The reason for the similarity of the results is apparent from panel (c) which compares labor supplies in efficiency units with hours worked in each (school, year) cell. Even though the model implies substantial changes in the relative efficiencies of college graduates versus high school graduates, changes in labor supply are dominated by changes in hours worked. These are due to the expansion of education and to changes in the age-composition of the workforce over time.

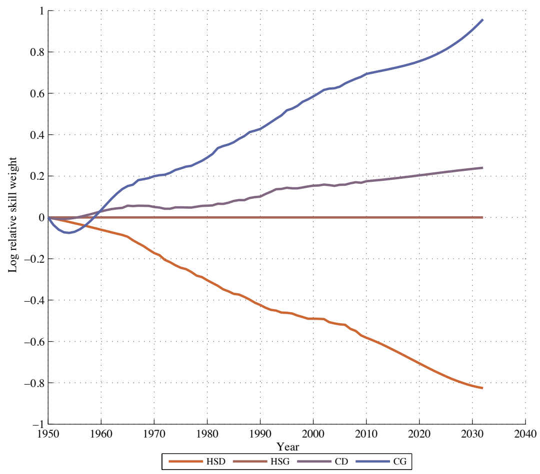
The elasticities of substitution between skill types are far larger than typical estimates suggest. However, they are poorly identified. Fixing the substitution elasticity between college graduates and other workers at 1.6, which is close to the values estimated (Katz and Murphy, 1992) and more recently by Autor, Katz, and Kearney (2008), has minimal effects on the model predictions. However, it implies larger changes in relative labor efficiencies which are needed to offset the large increase in the hours worked by skilled relative to unskilled workers. Notably, the relative efficiency of college graduates to high school graduates is implied to rise 20-fold between 1950 and 2010.

5 Results

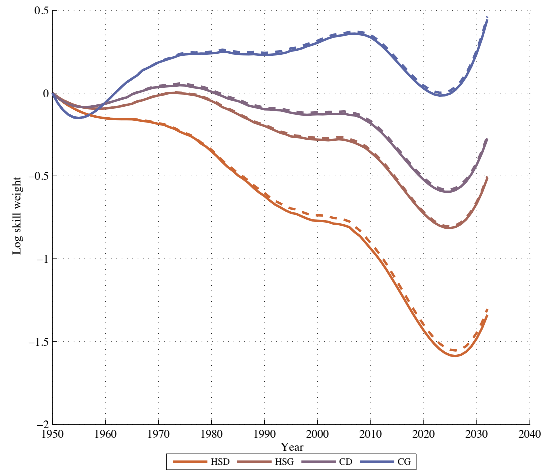
This section examines to what extent the Ben-Porath model can account for several measures of wage inequality that have been highlighted in the literature.

5.1 Overall Wage Dispersion

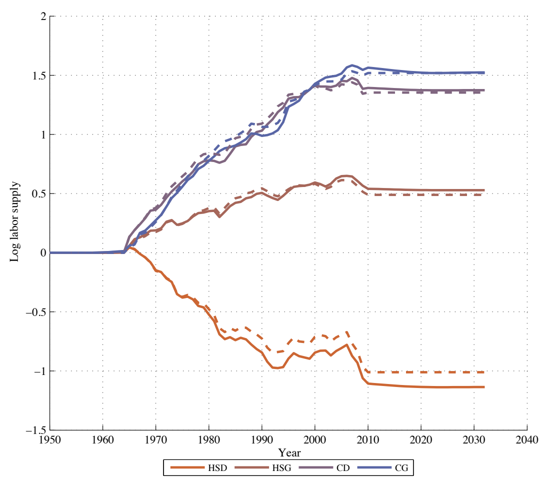
Figure 11 shows the evolution of cross-sectional wage inequality measures in the model and the data. Similar statistics are presented in Figure 3 of Autor, Katz, and Kearney (2008) and Figure 4 of Heathcote, Perri, and Violante (2010), even though both employ somewhat



(a) Skill weights relative to HSG: $\ln(\omega_{s,\tau}/\omega_{HSG,\tau})$



(b) Skill weights, $\ln(\omega_{s,\tau})$



(c) Labor supply

Figure 10: General equilibrium results

different sample selection criteria.[12]

The model slightly overstates the standard deviation of log wages in all years (panel a). However, it replicates the levels and changes of the 90/10 wage ratio (panel b), the 90/50 ratio (panel c) and the 50/10 ratio (panel d). The model also accounts for the fact that wage inequality was flat before 1970.

The model does not account for the fact that the 50/10 ratio grows only between 1970 and 1990 while the 90/50 ratio rises only after about 1980. In the model, the entire wage distribution spreads out roughly at the same time. The literature attributes the differential timing to changes in labor market institutions during the 1980s, which mostly affected low wage workers. The model abstracts from these changes.

Figure 12 shows that the standard deviation of log wages behaves very differently for different school groups. The most drastic increase occurs among college graduates after about 1990. The model matches this increase and also the approximate timing. Among high school dropouts, wage dispersion is roughly flat. High school graduates and college dropouts exhibit gradually increasing wage dispersion, which the model does not fully replicate.

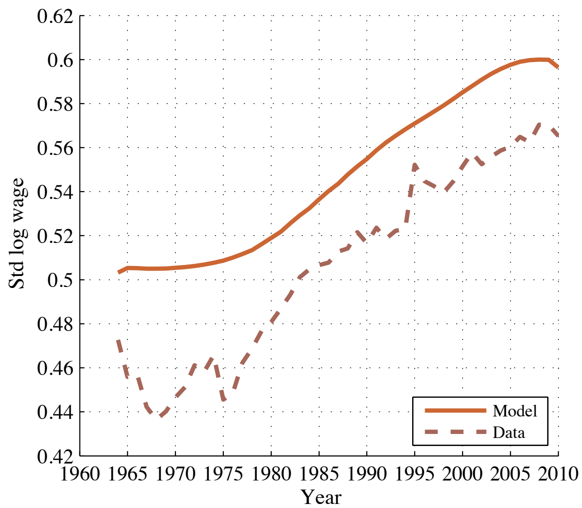
Fanning out of the wage distribution. Figure 13 shows the change in log wages experienced by each wage percentile over the period $1964-2010$.[13] The model replicates the fact that higher wage percentiles experienced stronger wage growth. The main discrepancy occurs at the very bottom of the wage distribution.

5.2 College Wage Premium

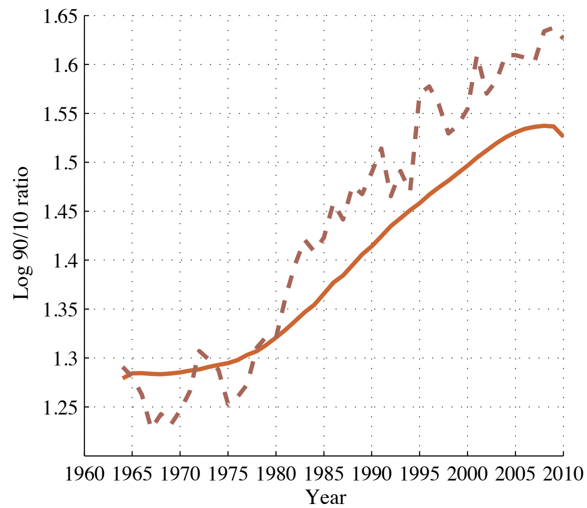
Figure 14 shows the evolution of the college wage premium. This is computed as a constant-composition average of mean log wages of college graduates relative to high school graduates.

[12] Autor, Katz, and Kearney (2008) focus on full-time, full-year workers. Heathcote, Perri, and Violante (2010) include all men who work at least 260 hours per year. The papers also differ in their handling of extreme wage observations.

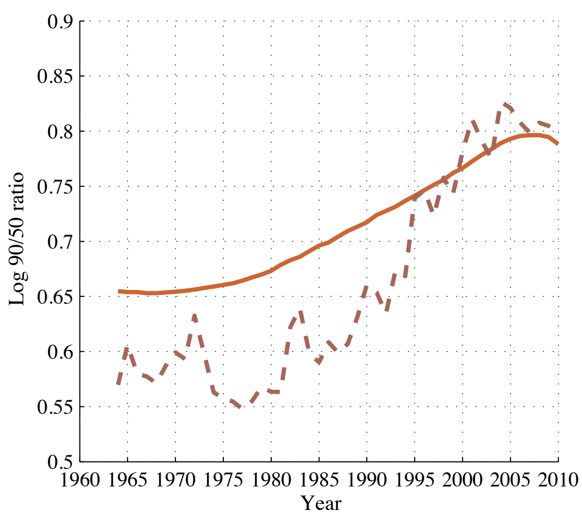
[13] This is similar to Figure 1 in Autor, Katz, and Kearney (2008) with two exceptions. First, I use composition adjusted wages, which removes the effects of rising schooling and experience. As a result, all wages changes are about 25% higher than those reported by Autor, Katz, and Kearney (2008). Second, my data show a smaller gap between high and low wage percentiles. In large part this is due to the fact that Autor, Katz, and Kearney (2008) restrict their sample to full-time/full-year workers.



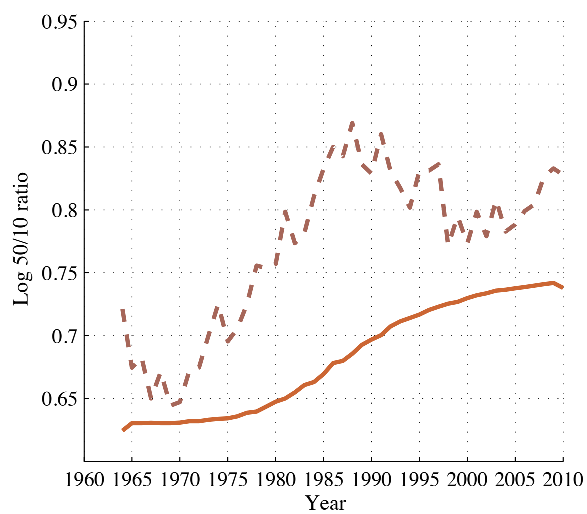
(a) Standard deviation of log wages



(b) Log 90/10 wage ratio

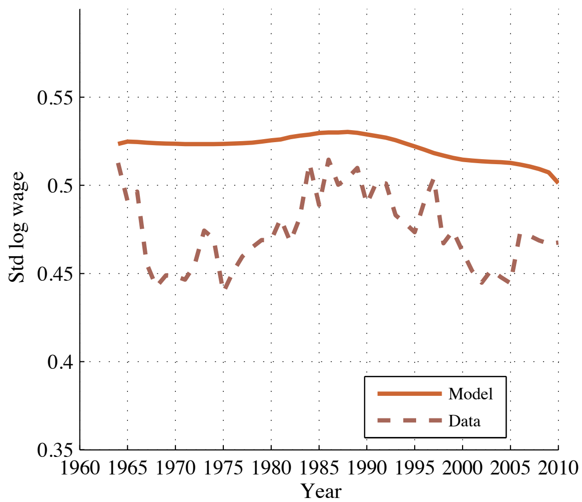


(c) Log 90/50 wage ratio

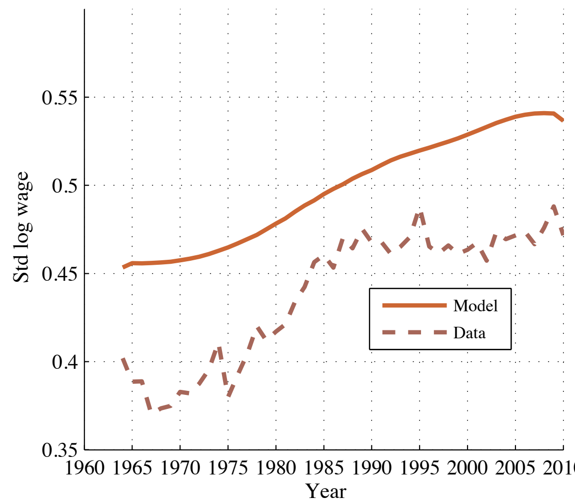


(d) Log 50/10 wage ratio

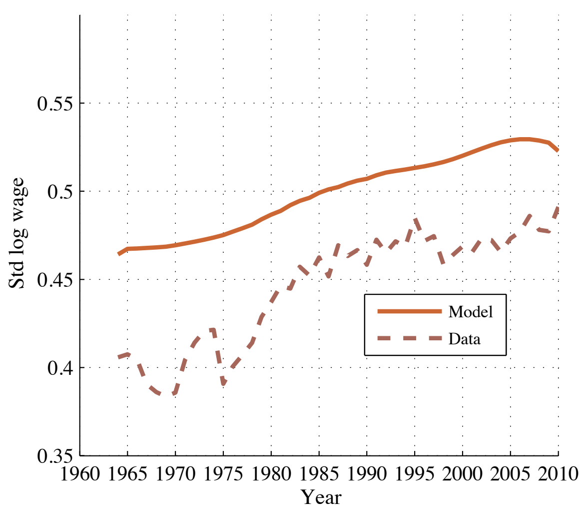
Figure 11: Evolution of U.S. Wage Inequality



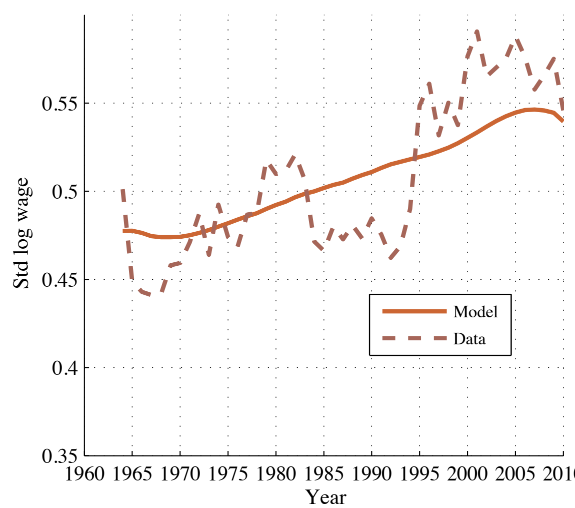
(a) High school dropouts



(b) High school graduates



(c) College dropouts



(d) College graduates

Figure 12: Evolution of U.S. Wage Inequality

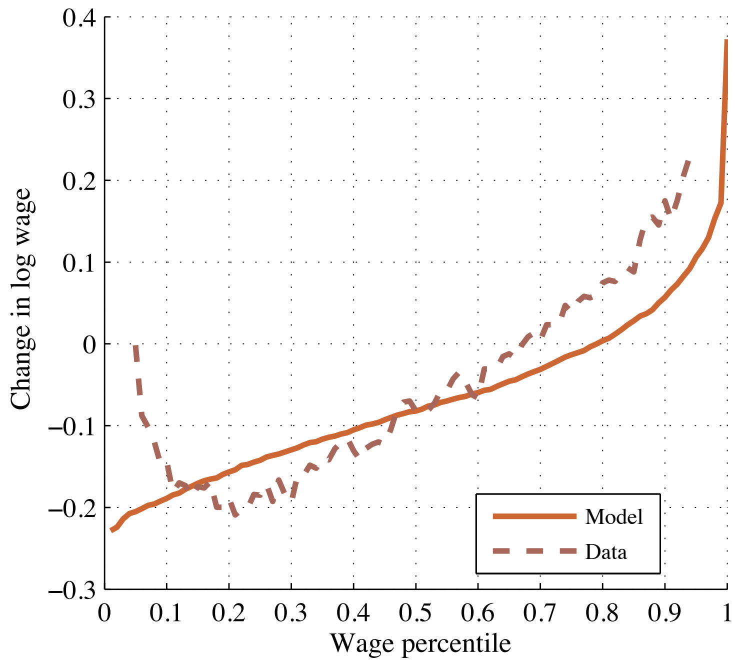


Figure 13: Evolution of U.S. Wage Inequality

Notes: The figure shows the change in log wages experienced by each wage percentile over the period $1964 - 2010$.

The model replicates the overall rise in the college premium as well as its initial decline before 1980. It also accounts for the differential timing and magnitude for young versus old workers. The mechanism is, of course, entirely different from the one proposed by Card and Lemieux (2001) who propose a model where young and old workers are imperfect substitutes.

How does the model account for the differential evolution of the young and old skill premiums? Figure 15 examines the evolution of the young college premium. Until 1975, young high school graduates increase their training time, while college graduates do not. As a result, human capital and labor efficiency rise among high school graduates, but not among college graduates. This overrides the increase in the relative skill prices of college graduates over this period, so that the college wage premium declines. After 1975, training time and human capital of high school graduates and college graduates move in parallel until about 1990. The rising college skill price then causes the college wage premium to grow.

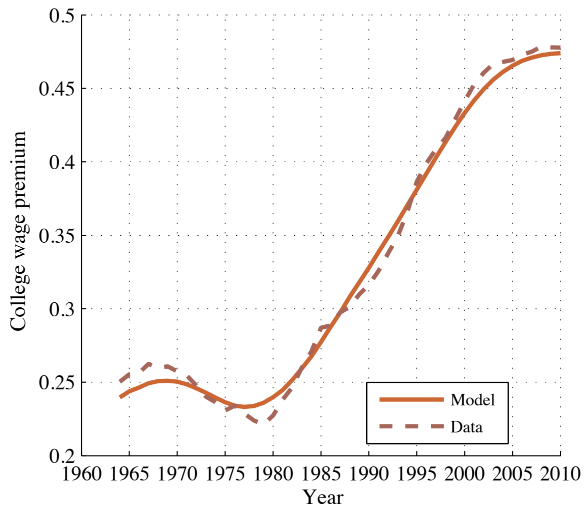
Among older workers (not shown), changes in human capital investment are much smaller, owing to their shorter time horizon. As a result, the college wage premium is largely driven by relative skill prices, which start to increase in the early 1970s.

5.3 Returns to Experience

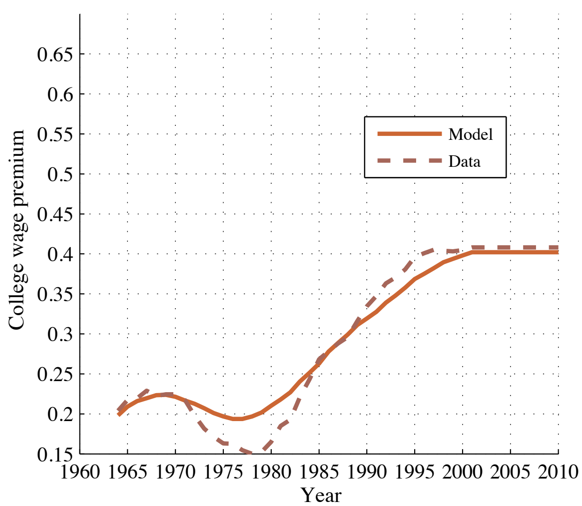
Based on cross-sectional data, studies concluded that returns to experience increased over time (Heathcote, Perri, and Violante, 2010; Eckstein and Nagypal, 2004). Kambourov and Manovskii (2009) point out an alternative interpretation: the longitudinal wage profiles of cohorts born after 1942 get flatter over time. Figure 16 follows their approach and shows the change of mean log wages between the ages $25 - 40$.

For all school groups, returns to experience decline until roughly the 1950 cohort and rise thereafter. The model replicates these patterns. The main discrepancy occurs of for college graduates, where the model age profiles are flatter than the data profiles for the early cohorts.

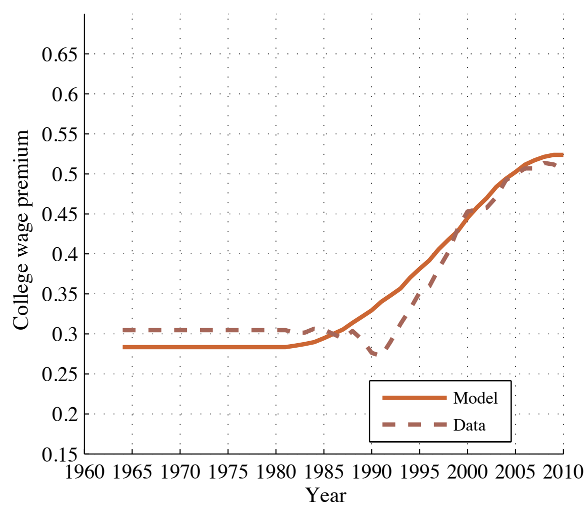
In the model, the slope of each cohort's wage profile is largely determined by growth of work time. Figure 17 shows that the growth rates of $(\ell - l)/\ell$ over the life-cycle are u-shaped for all school groups. The troughs occur near the troughs of the wage growth rates shown in Figure 16. Since compositional variation across cohorts is small, variation in training time



(a) College wage premium

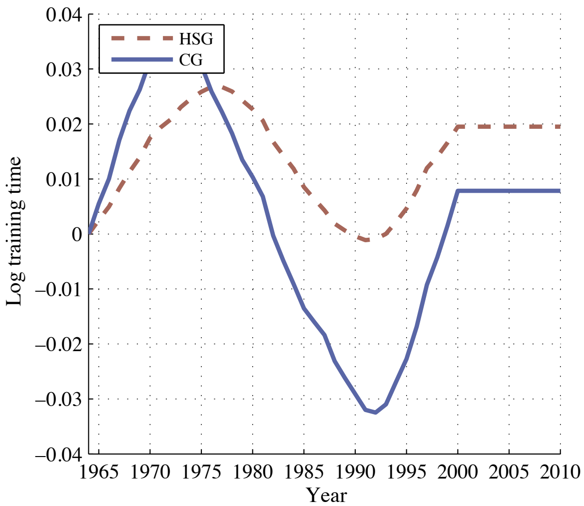


(b) College wage premium. Young workers

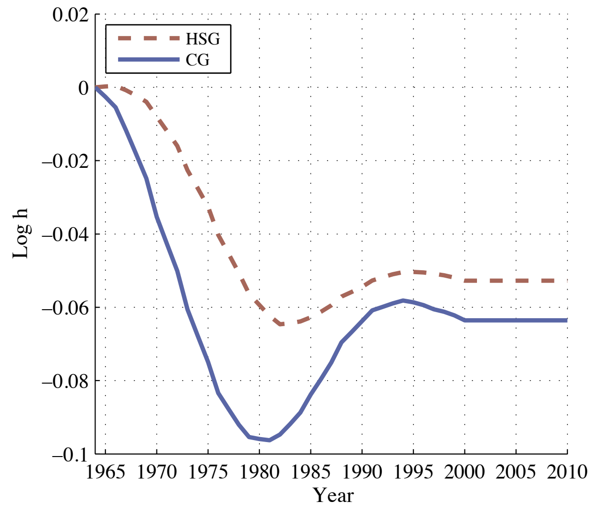


(c) College wage premium. Older workers

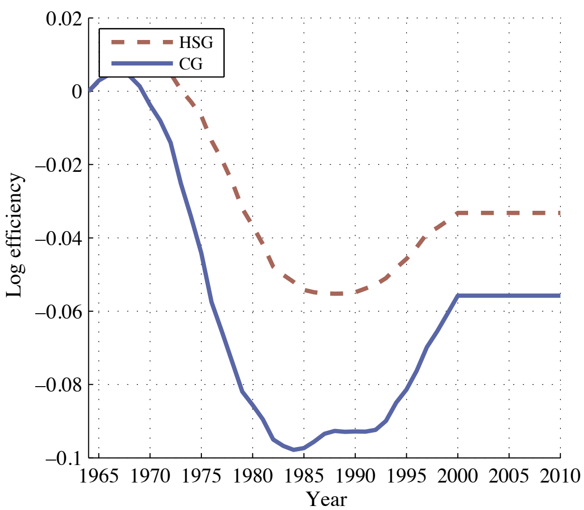
Figure 14: College wage premium



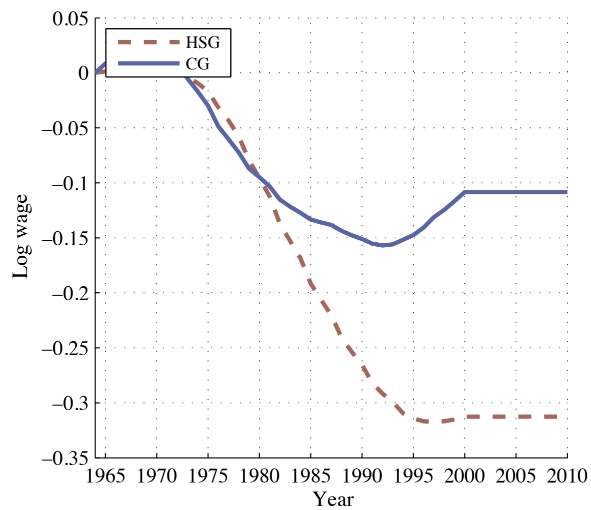
(a) Training time



(b) Human capital

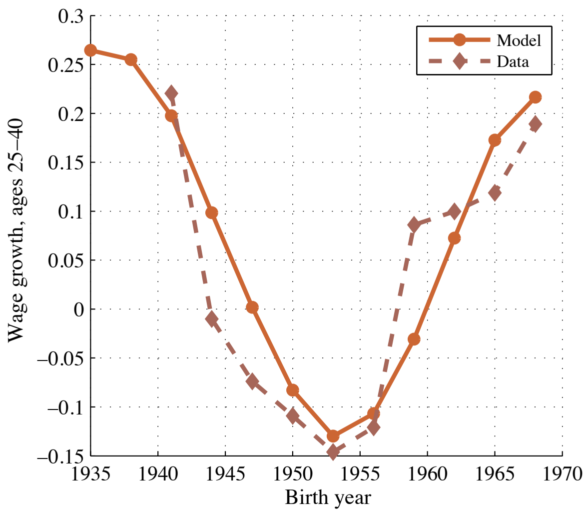


(c) Labor efficiency, $h(\ell - l)/\ell$

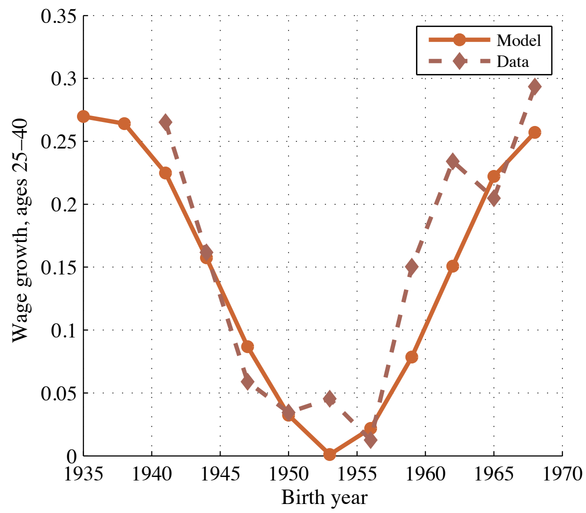


(d) Wages

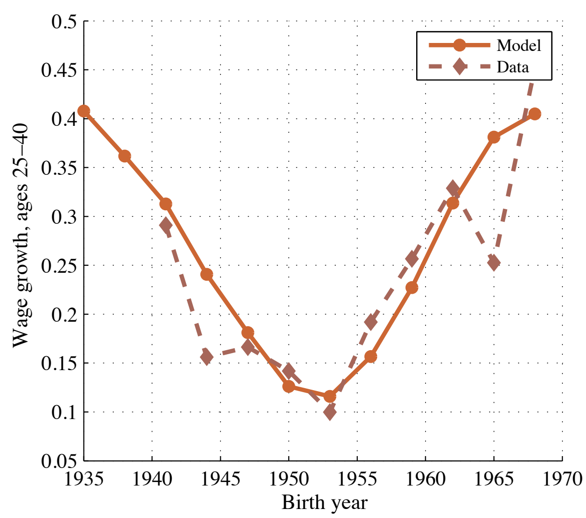
Figure 15: Young workers



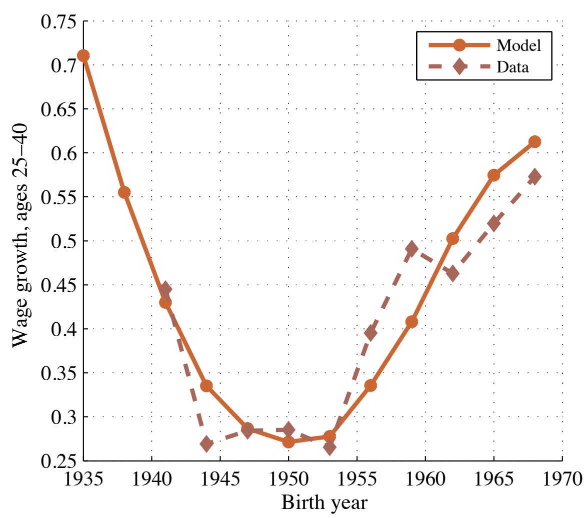
(a) High school dropouts



(b) High school graduates



(c) College dropouts



(d) College graduates

Figure 16: Wage growth over the life-cycle

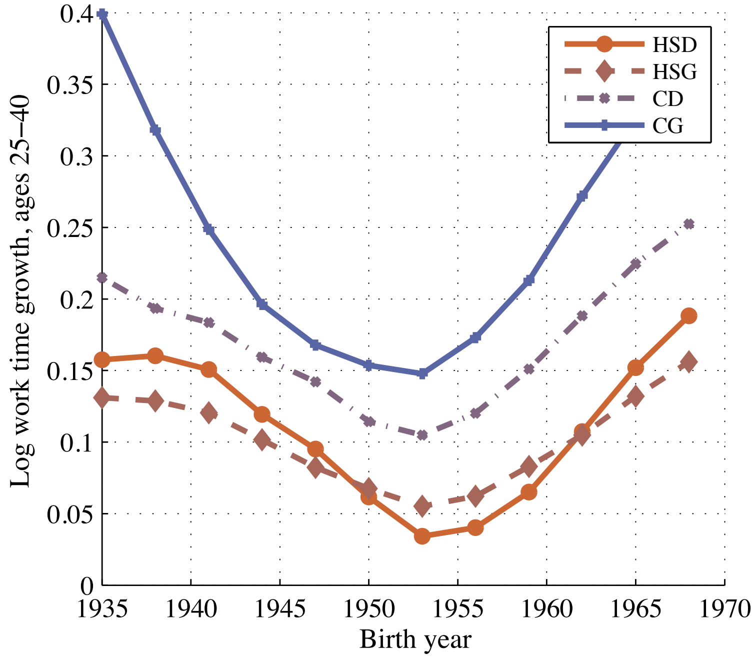


Figure 17: Growth of work time, $(\ell - l)/\ell$

is largely determined by future skill price growth rates (see equation 2). This mechanism is very different from Kambourov and Manovskii (2009) who highlight changes in occupational mobility.

5.4 Residual wage dispersion

Juhn, Murphy, and Pierce (1993) argue that roughly half of the rise in overall wage inequality occurs within age-school groups. Figure 18 shows the evolution of residual wage inequality in the model and in CPS data. Residual log wages are obtained from cross-sectional regressions of log wages on school dummies, and school dummies interacted with cubic polynomials in experience. In the data, I also include dummies for being married, white race, and region of residence. The regression uses constant composition weighted least squares.

In the data, the residual standard deviation rises from 0.5 to 0.58, starting in 1975. The model replicates this increase as well as the approximate timing. The counterfactual results reported in Section 5.5 show that almost the entire increase is due to the rising variance of persistent wage shocks which was documented by Gottschalk and Moffitt (1994).

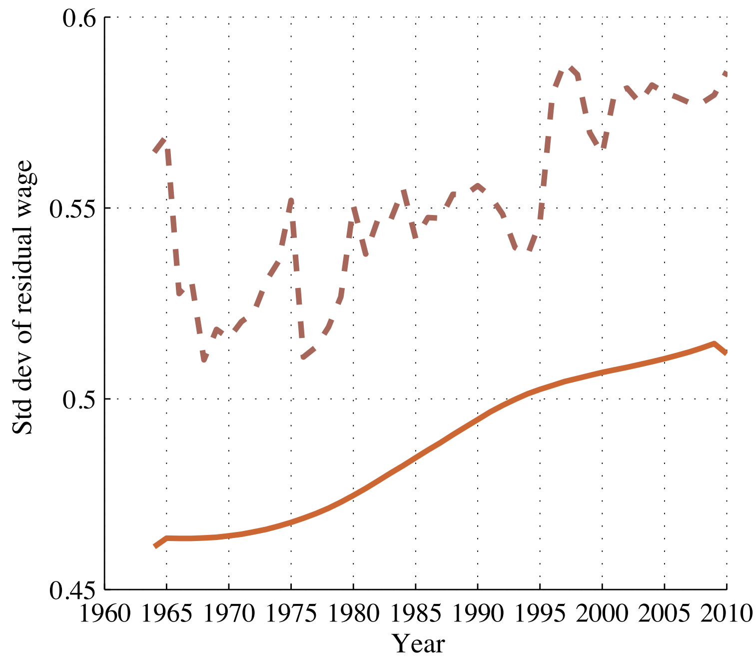


Figure 18: Residual standard deviation of log wages

5.5 Counterfactual Experiments

In order to gain insight into the forces that give rise to increasing wage inequality over time, I consider a number of counterfactual experiments. Each experiments fixes one variable relative to the baseline model. The agents' schooling and human capital investment decisions are held fixed, so that the experiments capture only the direct effects of the changes. The experiments are:

1. Fixed skill prices: $w_{s,\tau}$ is held fixed at $w_{s,1964}$. This captures the mechanical effect of changing (relative) skill prices on the wage distribution (what Guvenen and Kuruscu (2009) call the price effect).

2. Fixed h: Human capital investment and thus h are fixed for each agent at the values taken for the first model cohort. This captures the effects of changing human capital investments induced by changing skill prices and worker endowments.

3. Fixed shock variances: $\sigma_{\xi,s,\tau} = \sigma_{\xi,s,1964}$.

4. Fixed schooling: Each worker's schooling decisions are fixed at the values for the first

model cohort. This captures the effects that rising schooling has on the endowments of workers in different school groups.

For each experiment, I recompute wage histories and aggregate wage distribution statistics (using constant-composition weights). Table 4 summarizes the results. It shows the changes over the sample period of several indicators. The implications are as follows:

1. The rise in overall wage inequality as measured by the standard deviation of log wages, the 90/50 or the 50/10 wage ratio is due in roughly equal parts to diverging skill prices and larger shocks.

   Roughly the same is true for the increase in lifetime earnings dispersion.

2. Changes of within group dispersion, such as the standard deviation of residual log wages or of wages within school groups, are mostly due to rising shock variances.

3. The rising college wage premium is almost entirely the mechanical implications of diverging skill prices.

4. The changing returns to experience mostly stem from time-varying skill prices which are magnified by changing human capital investments.

5. The expansion of education is not a major contributor to any of the changes (except through its effect on skill prices).

For completeness, the last two columns of Table 4 show the skill price and shock variance experiments where agents update their human capital decision rules in response to the changed incentives. Aside from the effect of skill prices on the returns to experience, the indirect effects of the shocks affecting wage inequality are generally small. While the expansion of schooling and time variation in skill prices induce variation in human capital investments, these have little impact on within group inequality or relative wages. This finding supports the validity of reduced form or accounting decompositions of changing wage inequality which abstract from indirect effects (e.g., Juhn, Murphy, and Pierce 1993). The remainder of this Section explores why the indirect effects are so small.

		h fixed				h updated	
	Baseline	w	Schooling	h	Shocks	w	Shocks
Std log wage	0.09	0.45	-0.02	0.03	0.48	0.39	0.52
90/50 ratio	0.13	0.47	-0.11	0.02	0.56	0.43	0.58
50/10 ratio	0.11	0.46	0.03	0.03	0.43	0.38	0.49
Std log lifetime earnings	0.09	0.52	-0.03	-0.04	0.38	0.47	0.39
Std residual log wage	0.05	0.01	-0.03	-0.06	1.07	-0.13	1.11
Std log wage HSG	0.08	0.01	-0.01	-0.03	1.03	-0.10	1.08
Std log wage CG	0.06	0.06	-0.07	0.08	0.87	-0.01	0.98
College premium	0.23	0.96	-0.01	0.03	-0.01	1.00	-0.02
Cohort wage growth HSG	0.27	0.65	0.02	0.33	-0.02	0.84	-0.04
Cohort wage growth CG	0.44	0.41	0.08	0.60	0.01	0.83	-0.01

Table 4: Counterfactual experiments

Notes: The “Baseline” column shows the changes of several wage dispersion measures over the period 1964 – 2010. For each counterfactual experiment, the table shows the fraction of the “baseline” change that is due to the variable held constant in the experiment. Experiment w sets $w_{s,\tau} = w_{s,1964}$. “Schooling” fixes school choices at the level of the first model cohort. Experiment h fixes human capital investment at the level of the first model cohort. “Shocks” fixes persistent wage shocks: $\sigma_{\xi,s,\tau} = \sigma_{\xi,s,1964}$. The last two columns show the w and shocks experiments, allowing agents to update their human capital investments.

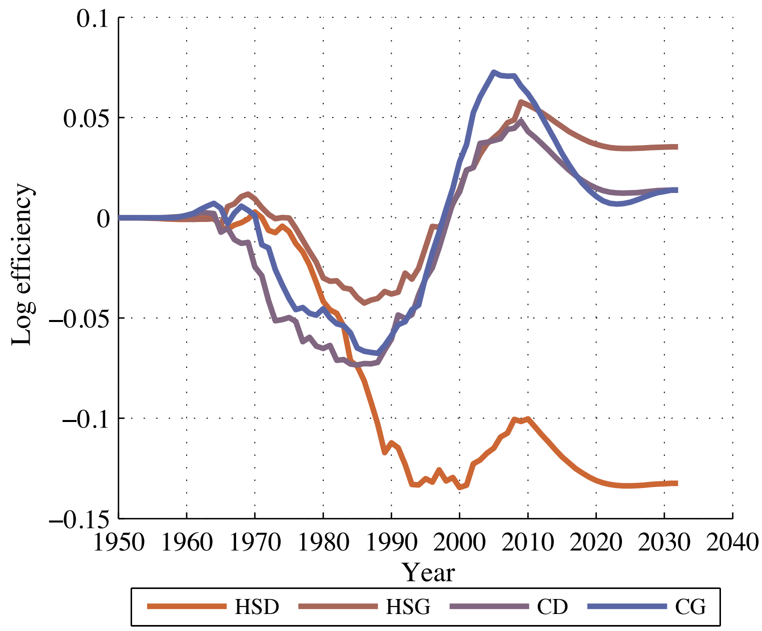


Figure 19: Labor efficiencies

College wage premium: The measured college wage premium changes for two reasons: (i) the relative skill price paid for college educated labor rises over time, and (ii) changing skill prices induce time variation in human capital investment and labor efficiency. The latter, indirect effect is quite small (see Table 4).

The reason can be seen in Figure 19 which shows the time paths of labor supplies per hour worked (the mean of e) in each school group. Labor efficiencies change substantially over time, rising by about 10 percentage points between their troughs in 1990 to their peaks around 2005. However, the magnitudes and timing of the changes are similar across school groups (except for high school dropouts) and therefore largely cancel in their effects on skill premiums.

Within group inequality: The changes in within group inequality are almost entirely due to time variation in shock variances. Changes in human capital investment make almost no contribution, even though they are far from negligible (Figure 19). Inspecting the optimal investment condition (10), which is repeated here for convenience, reveals one reason for this result:

$$(h_{i,s,c,t}l_{i,s,c,t})^{1-\alpha_s} = \frac{\alpha A(a,s)}{(1-\delta_s)} \sum_{j=1}^{T-t} X_{s,c,t,j} \frac{\mathbb{E}(q_{i,s,c,t+j}|t)}{q_{i,s,c,t}} \tag{19}$$

Changing future skill prices affect the values of $X_{s,c,t,j}$. Aside from the interaction with the expected growth of the productivity shock q, the response of human capital investment to wage shocks is the same for all persons in a given school/age group.

Conversely, changing shock variances affect human capital investment through the expected growth of the productivity shock q in (19).[14] Yet this does not magnify the effect on within group inequality because the human capital investment response is the same for all agents.

These results contrasts with the model of Guvenen and Kuruscu (2009), where a key part of the argument relies on the notion that a one time increase in the a skill price increases the human capital investment of highly able workers more than that of less able workers. The result is a compression of the within group wage distribution in the short run, followed by an increase in wage inequality later on as the additional human capital begins to pay off.

The expansion of educational attainment: An increase in cohort education affects the means and variances of abilities and human capital endowments. Mean abilities decline in all school groups as less able students attain more advanced degrees. The changes in mean human capital and endowment variances are theoretically ambiguous (see the discussion in Acemoglu 2002a and Hendricks and Schoellman 2011). However, the model implies that these changes in endowment distributions have negligible effects on all of the inequality measures shown in Table 4.

One reason is that the model implies essentially no correlation between school choice and human capital endowments. Initial human capital has two opposing effects on school choice that approximately cancel each other out. Holding ability fixed, agents with higher human capital endowments wish to invest less in human capital and therefore leave schooling earlier. On the other hand, human capital and abilities are positively correlated and higher abilities induce more human capital investment and longer schooling. In the calibrated model the net effect is that the distribution of human capital in each school group changes little over time.

[14] Transitory shock variances do not affect household decisions at all.

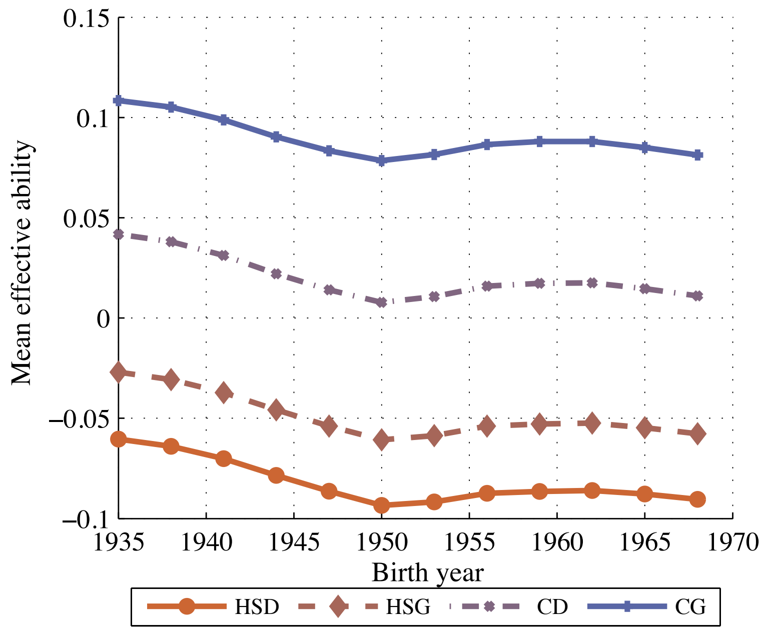


Figure 20: Mean ability by schooling and cohort

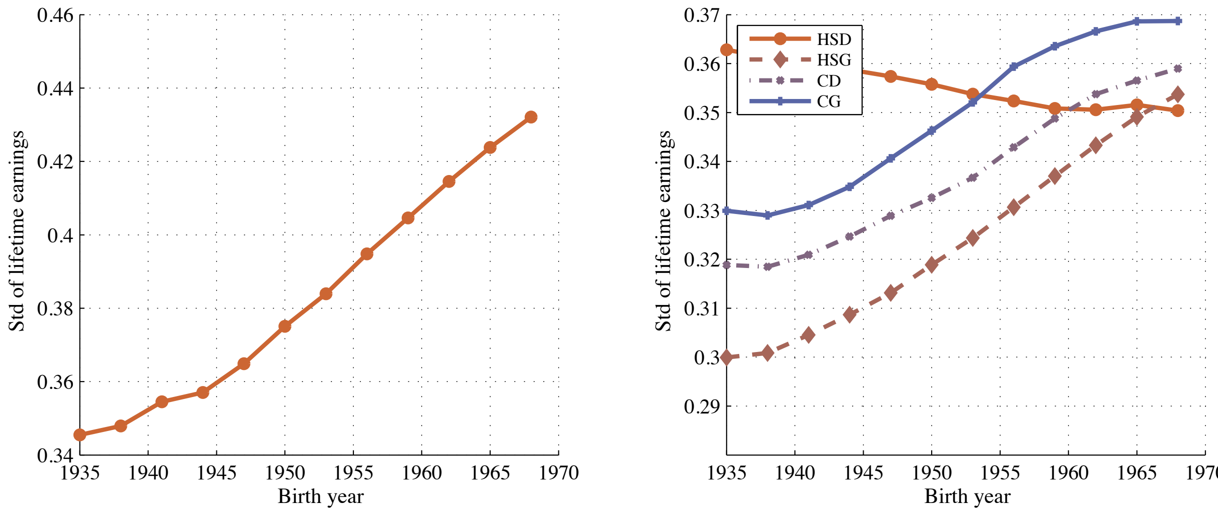
The model does imply substantial school sorting by ability. Figure 20 shows the evolution of mean abilities by schooling. The mean ability of college graduates is about 1 standard deviation above that of high school graduates. As expected, mean abilities decline as schooling expands, roughly until the 1950 birth cohort. However, the mean abilities of all school groups decline by similar amounts (0.02 to 0.03). The model also generates little change in standard deviation of abilities given schooling. As a result, the expansion of education has little effect on wage dispersion between or within school groups.

5.6 Lifetime Earnings

One benefit of working with a structural model is that the implications for lifetime inequality can be computed. Given that households are forward looking, lifetime inequality matters for welfare more than period inequality.

Figure 21 shows the standard deviation of log lifetime earnings for all model cohorts.[15] It rises by 8 percentage points, which is of comparable magnitude with the rise in the

[15] As in all aggregate statistics, composition changes are removed by weighting each school group with its average mass across years.



(a) Pooled school groups

(b) School groups

Figure 21: Standard deviation of lifetime earnings

cross-sectional standard deviation of wages. Most of the increases occur among high school graduates and college graduates, reflecting the fact that their shock variances rise over time (see Table 2).

This result differs from Guvenen and Kuruscu (2009). In their model the dispersion of lifetime earnings is essentially constant over time. One reason is that Guvenen and Kuruscu abstract from rising shock variances. As pointed out earlier, their model accounts for rising within group inequality through the differential human capital investment behavior of high and low ability agents, which is not present in my model.

Predictability. Figure 22 examines the predictability of lifetime earnings. Following Huggett, Ventura, and Yaron (2011), this is defined as the variance of expected lifetime earnings given the agent's information set at the beginning of work to the ratio of realized lifetime earnings. In contrast to the findings of Huggett, Ventura, and Yaron (2011) predictability is generally quite low at about 0.25 for high school educated workers and 0.1 for college educated workers. Pooling school groups, as is implicitly done by Huggett et al., yields a predictability ratio between 0.2 and 0.25. Predictability declines over time for high school graduates and college graduates, again reflecting the fact that shock variances rise over time (see Table 2).

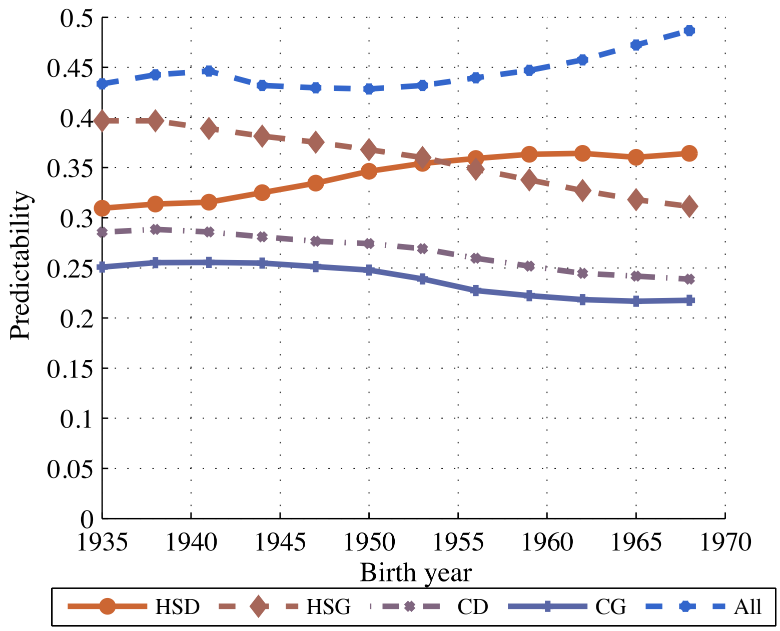


Figure 22: Predictability of lifetime earnings

Given that my model shares many features with Huggett et al.'s, the difference in results is intriguing. To gain some intuition, it is helpful to calculate the asymptotic variance of human capital as agents approach their steady states. Given that $l_t \approx 0$ for older workers, this is also the variance of wages net of shocks. The law of motion (2) together with the optimal investment condition (10) imply that $h_{t+1} = (1 - \delta_s) h_t + A (C_{1t}A)^{\alpha_s/1-\alpha_s}$, where $C_{1t} > 0$ is a function of parameters. In steady state $h_{t+1} = h_t = h_{ss}$, so that $\ln h_{ss} = C_2 + \ln (A) / (1 - \alpha)$, where C_2 is a constant. It follows that

$$Var (\ln h_{ss}) = \left(\frac{\theta}{1 - \alpha_s} \right)^2 \tag{20}$$

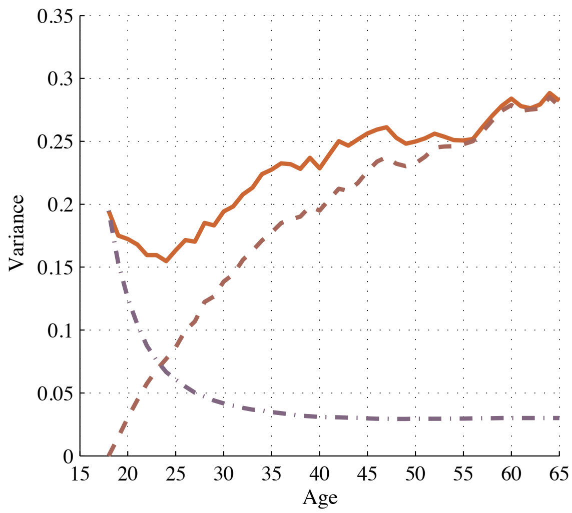
The variance of steady state human capital equals the variance of effective ability (θ^2) scaled by the curvature of the job-training technology. The calibrated parameters of θ near 0.1 and α_s near 0.5 imply a steady state variance near 0.04 (where I neglect to correct for the fact that the variance of abilities conditional on schooling is somewhat smaller than 1).

Figure 23 illustrates how this logic plays out over the workers' life-cycles. For each school group, the figure shows the variance of log wages by age for the cohort born in 1950. It also decomposes this variance into the contribution of shocks ($Var (\ln q_t)$) and of effective labor input ($Var (\ln [h_t (\ell_t - l_t)])$). The variance of effective labor inputs declines with age and approaches a value that is consistent with the steady state calculations. Given that the variance of log wages rises with age towards values near 0.35, the bulk of the wage dispersion among older workers is due to shocks.

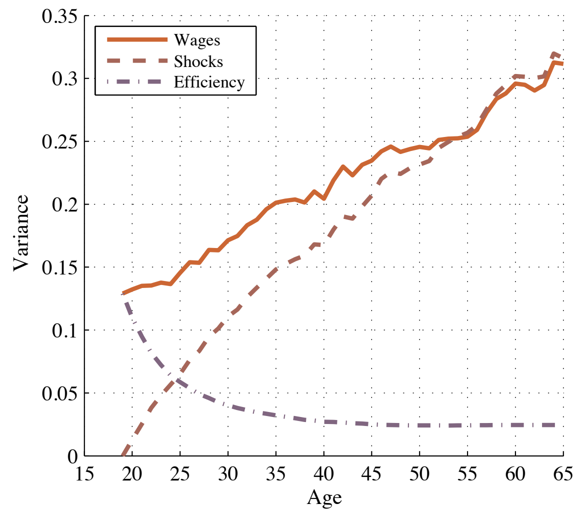
Since the variances of h_{ss} in Huggett et al.'s model obeys (20), it is possible to isolate one source of higher predicability in their model. While their value of $\theta = 0.12$ is close to the one calibrated here, Huggett et al.'s estimate of $\alpha = 0.7$ implies $Var\left(\ln h_{ss}\right) \approx 0.13$ versus 0.04 in my model. The more linear technology magnifies the long-term effects of ability differences. Of course, Huggett et al.'s analysis differs in several respects from mine. A full reconciliation of the different results is left for future research.

6 Conclusion

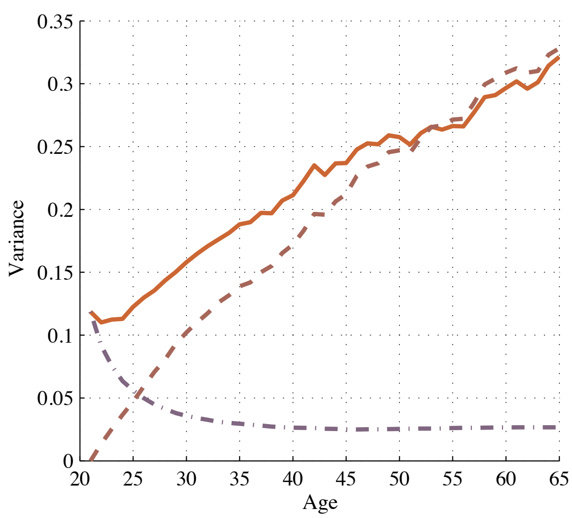
To be written.



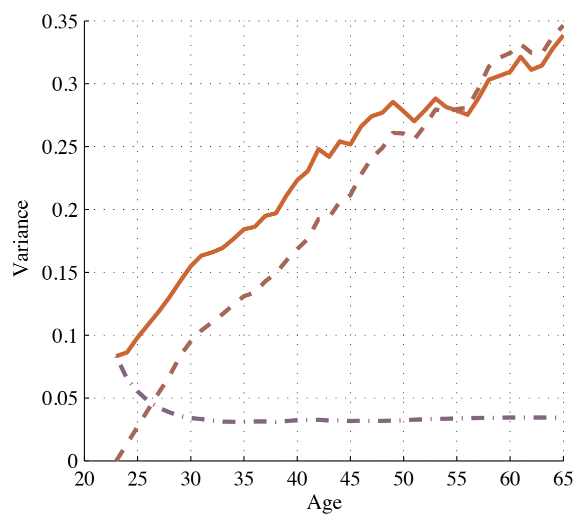
(a) High school dropouts



(b) High school graduates



(c) College dropouts



(d) College graduates

Figure 23: Wage dispersion by age

References

Acemoglu, D. (2002a): "Directed technical change," *Review of Economic Studies*, 69, 781–810. 43

——— (2002b): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, pp. 1–67. 2

Altonji, J. G., and C. R. Pierret (2001): "Employer Learning and Statistical Discrimination," *The Quarterly Journal of Economics*, 116(1), pp. 313–350. 16

Autor, D. H., L. F. Katz, and M. S. Kearney (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics*, 90(2), 300–323. 2, 5, 28, 30, 54

Berman, E., J. Bound, and Z. Griliches (1994): "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufacturers," *The Quarterly Journal of Economics*, 109(2), 367–397. 2

Berman, E., J. Bound, and S. Machin (1998): "Implications of Skill-Biased Technological Change: International Evidence," *The Quarterly Journal of Economics*, 113(4), 1245–1279. 2

Bound, J., and G. Johnson (1992): "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations," *The American Economic Review*, 82(3), pp. 371–392. 2

Bowlus, A. J., and C. Robinson (2012): "Human Capital Prices, Productivity, and Growth," *The American Economic Review*, 102(7), 3483–3515. 14, 54

Bureau of Labor Statistics; US Department of Labor (2002): "National Longitudinal Survey of Youth 1979 cohort," Produced and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH. 59

Card, D., and J. DiNardo (2002): "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20(4), 4. 5

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *The Quarterly Journal of Economics*, 116(2), pp. 705–746. 2, 5, 34

Caselli, F. (2009): "Comment," *NBER Macroeconomics Annual*, 24(1), pp. 301–314. 12

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5), pp. 1001–1044. 5

Eckstein, Z., and E. Nagypal (2004): "The evolution of US earnings inequality: 1961-2002," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(2), 10–29. 2, 34

Goldin, C., and L. F. Katz (2008): *The Race between Education and Technology*. Harvard University Press. 55

Gottschalk, P., and R. Moffitt (1994): "The growth of earnings instability in the US labor market," *Brookings Papers on Economic Activity*, (2), 217–272. 2, 4, 38

Greene, W. (2012): *Econometric Analysis*. Prentice Hall, 7 edn. 11

Guvenen, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *The American Economic Review*, 97(3), pp. 687–712. 24, 58

Guvenen, F., and B. Kuruscu (2009): "A Quantitative Analysis of the Evolution of the US Wage Distribution: 1970-2000," *NBER Macroeconomics Annual*, 24(1), 227–276. 6, 7, 12, 39, 43, 45

Guvenen, F., B. Kuruscu, and S. Ozkan (2011): "Taxation of human capital and wage inequality: A cross-country analysis," Mimeo. University of Minnesota. 7

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967-2006," *Review of Economic Dynamics*, 13(1), 15 – 51, Special issue: Cross-Sectional Facts for Macroeconomists. 2, 28, 30, 34

Heathcote, J., K. Storesletten, and G. L. Violante (2010): "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of political economy*, 118(4), 681–722. 7

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), 1–58. 6, 7, 19

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. Hanushek, and F. Welch, vol. 1, chap. 7, pp. 307–458. Elsevier. 11

Hendricks, L. (2012): "The Ben-Porath Model and Age-wage Profiles," Manuscript. University of North Carolina. 19

——— (2013): "Identifying the Technology of Human Capital Production," Manuscript. University of North Carolina. 53, 58

Hendricks, L., and T. Schoellman (2011): "Student Abilities During the Expansion of U.S. Education, 1950-2000," Mimeo. University of North Carolina. 7, 11, 16, 18, 43

Huggett, M., G. Ventura, and A. Yaron (2006): "Human Capital and Earnings Distribution Dynamics," *Journal of Monetary Economics*, 53(2), 265–290. 14

——— (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101, 2923–54. 6, 7, 45

Jaeger, D. A. (1997): "Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers," *Journal of Business and Economic Statistics*, 15(3), pp. 300–309. 54

Juhn, C., K. Murphy, and B. Pierce (1993): "Wage Inequality and the Rise in the Return to Skill," *Journal of Political Economy*, 101(3), 410–442. 38, 40

Kambourov, G., and I. Manovskii (2009): "Accounting for the Changing Life-Cycle Profile of Earnings," *University of Pennsylvania. Mimeo.* 2, 34, 38

Katz, L. F., and D. H. Autor (1999): "Changes in the Wage Structure and Earnings Inequality," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3, pp. 1463–1555. Elsevier. 2

Katz, L. F., and K. M. Murphy (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107(1), 35–78. 2, 4, 5, 13, 19, 24, 28

King, M., S. Ruggles, J. T. Alexander, S. Flood, K. Genadek, M. B. Schroeder, B. Trampe, and R. Vick (2010): "Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]," Minneapolis: University of Minnesota. 13, 53

Kuruscu, B. (2006): "Training and lifetime income," *The American Economic Review*, 96(3), 832–846. 19

Laitner, J. (2000): "Earnings within Education Groups and Overall Productivity Growth," *The Journal of Political Economy*, 108(4), pp. 807–832. 7

Lee, D., and K. I. Wolpin (2010): "Accounting for wage and employment changes in the US from 1968–2000: A dynamic model of labor market equilibrium," *Journal of Econometrics*, 156(1), 68–85. 7

Lemieux, T. (2007): "The changing nature of wage inequality," *Journal of Population Economics*, 21(1), 21–48. 2

McGrattan, E. R., and E. C. Prescott (2000): "Is the Stock Market Overvalued?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 24(4). 16

Murphy, K., and F. Welch (1992): "The structure of wages," *The Quarterly Journal of Economics*, 107(1), 285–326. 2

Storesletten, K., C. Telmer, and A. Yaron (2004): "Concumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics*, 51(3), 609–33. 24

Taber, C. (2002): "Tax reform and human capital accumulation: Evidence from an empirical general equilibrium model of skill formation," *Advances in Economic Analysis & Policy*, 2(1), 1–36. 19

Taubman, P., and T. Wales (1972): *Mental ability and higher educational attainment in the 20th Century*. Berkeley: National Bureau of Economic Research and Carnegie Commission on Higher Education. 16, 19

Thornton, R., J. Rodgers, and M. Brookshire (1997): "On the interpretation of age-earnings profiles," *Journal of Labor Research*, 18(2), 351–365. 14

Table 5: Summary statistics for CPS data

Year	N	Avg N per cell	N range
1965	20106	162	24 - 372
1970	18600	150	34 - 279
1975	16702	135	40 - 264
1980	22892	185	55 - 378
1985	20703	167	48 - 371
1990	21896	177	47 - 466
1995	19770	159	42 - 318
2000	34568	279	51 - 528
2005	32528	262	39 - 439
2010	30235	244	38 - 473

Notes: N is the number of observations. Avg N per cell refers to the average number of observations in each (age, school) cell. N range shows the minimum and maximum number of observations in each cell. Cells cover age range 30-60.

A CPS Data

A.1 Sample

CPS data are obtained from King, Ruggles, Alexander, Flood, Genadek, Schroeder, Trampe, and Vick (2010). Summary information is provided in Table 5. The data construction largely follows Hendricks (2013). The sample contains men between the ages of 18 and 65 observed in the $1964 - 2011$ waves of the March Current Population Survey. Persons are dropped if at least one of the following conditions are true:

1. weeks worked per year < 20
2. hours worked per week < 20
3. labor earnings are zero or missing
4. employment is either unpaid or in the armed forces.

A.2 Individual Variables

The construction of individual variables is based on Bowlus and Robinson (2012). As discussed in Jaeger (1997), the coding of schooling changes in 1991. I use the coding scheme proposed in his tables 2 and 7 to recode HIGRADE and EDUC99 into the highest degree completed and the highest grade completed.

Hours worked per year are defined as the product of hours worked last week and weeks worked last year. Weeks worked per year are intervalled until 1975. Each interval is assigned the average of weeks worked in years after 1975 in the same interval. Until 1975, hours worked per week are only available for the previous week (HRSWORK). I regress hours worked on years of schooling and a quadratic in experience to impute hours worked. After 1975, I use usual hours worked per week (UHRSWRK).

Income variables: Labor earnings are defined as the sum of wage and salary incomes (INCWAGE). Wages are defined as labor earnings divided by hours worked. Outliers with less than 5% or more than 100 times the median wage are dropped.

Income variables are top-coded. As discussed in Bowlus and Robinson (2012), the frequency of top-coding and the top-coded amounts vary substantially over time. In addition, top-coding flags contain obvious errors. In most years, fewer than 2% of labor earnings observations appear to be top-coded. Following Autor, Katz, and Kearney (2008), I multiply top-coded amounts by 1.5 in years before 1988. From 1996 onwards, top-coded amounts are set to the average of all values above to top code. I leave these value unchanged. Between 1988 and 1995 there is no clear way of identifying top-coded values in IPUMS data because INCWAGE is the sum of two variables with different top codes. In these years I leave top-coded values unchanged.

To avoid top-coding issues, I drop the top and bottom 2% of wage observations from the data and from the simulated model data in each year when computing wage statistics (such as mean log wages). Since Bowlus and Robinson (2012) find that allocated values have little effect on the constructed wage series, I do not exclude them. Dollar values are deflated using the Consumer Price Index (all items, U.S. city average, series Id: CUUR0000SA0; see bls.gov).

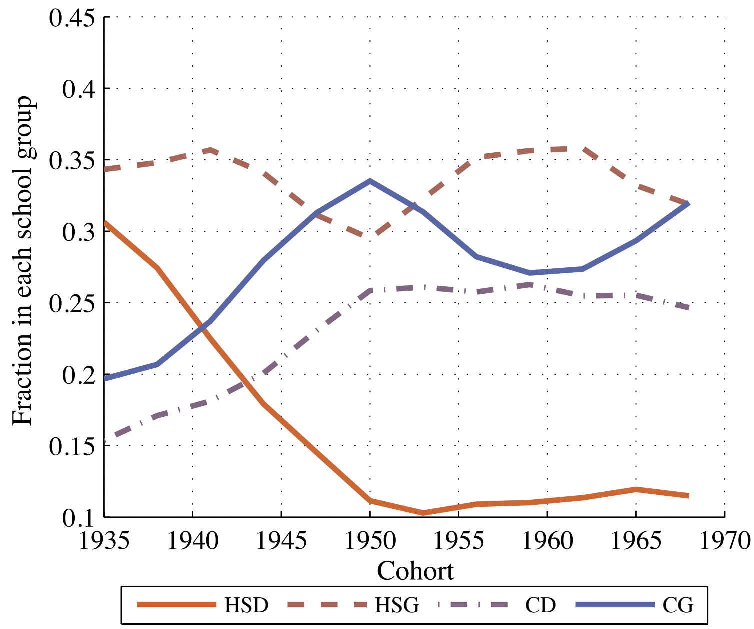


Figure 24: Educational Attainment By Birth Cohort

A.3 Cohort Variables

Schooling: The fraction of persons in cohort τ that achieves school level s is calculated by averaging over ages $30-40$ (not all ages are observed for all cohorts). Figure 24 shows these fractions. Each point represents one cohort. Educational attainment grows until the 1950 cohort and then levels off (see Goldin and Katz 2008 for an extensive discussion of these trends).

Age profiles of wage statistics: For each (age, school, cohort) cell, I compute the following statistics:

1. Mean and standard deviation of log wages among those reporting a valid wage.
2. Wages at various percentiles of the wage distribution in the cell.
3. Mass of persons in the cell.

The age profiles of the various statistics, such as mean log wages, are HP-filtered profiles computed from these cell statistics.

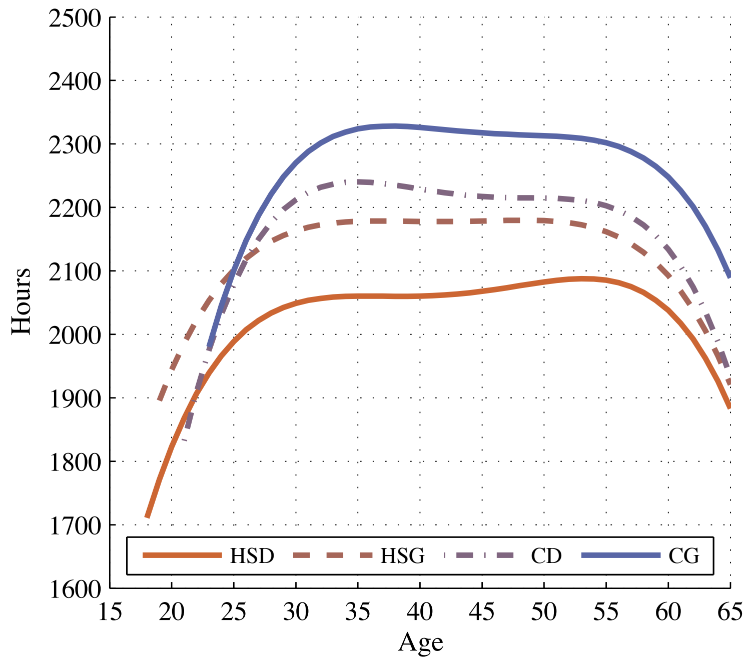


Figure 25: Age Hours Profiles

Age hours profiles: I construct the age profile of annual hours worked, $\ell_{s,t,\tau}$, as follows:

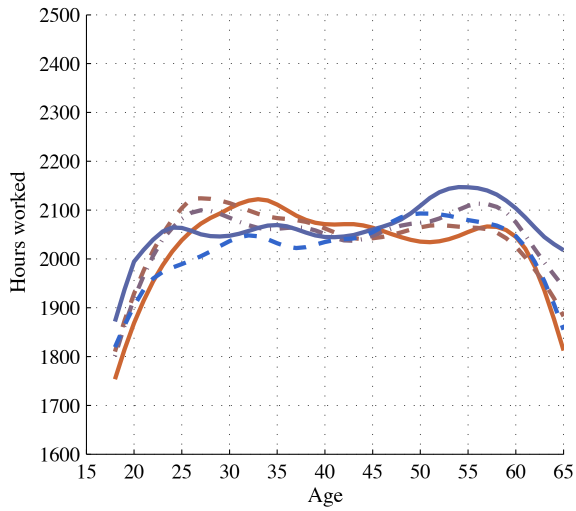
1. For combinations of s, t, τ that are observed, I set $\ell_{s,t,\tau}$ to the average hours worked of all persons in the cell divided by 2,000.

2. For each school group, I regress average hours worked in each s, t, τ cell on a quartic in age and on cohort dummies. Call the resulting hours values $\hat{\ell}_{s,t}$. Figure 25 shows these profiles.

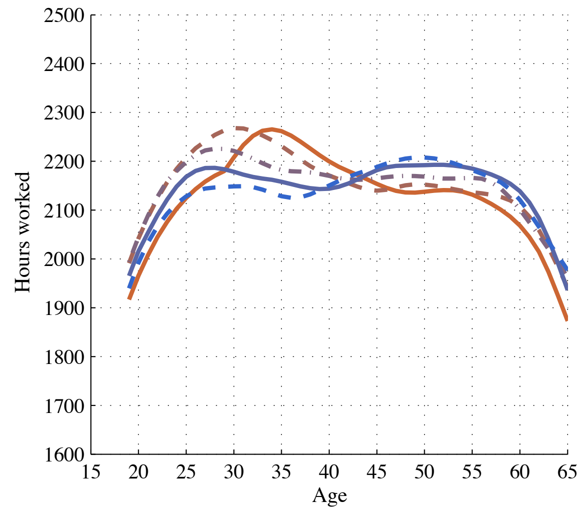
3. For combinations of s, t, τ that are not observed, I set $\ell_{s,t,\tau}$ equal to $\hat{\ell}_{s,t}$ which is scaled to match the level of $\ell_{s,t,\tau}$ for the nearest observed age.

4. Finally, the hours profile of each cohort is smoothed using an HP filter with parameter 20.

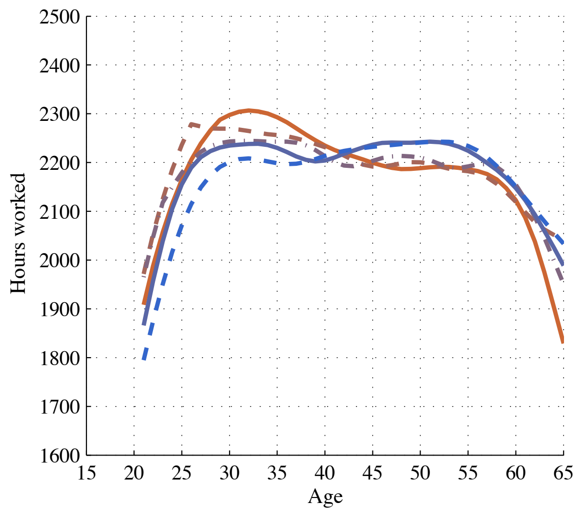
Figure 26 shows the age hours profiles for selected cohorts. For all school groups, the hours profiles get steeper over time. This implies that the incentives for investing in human capital around middle age get stronger over time.



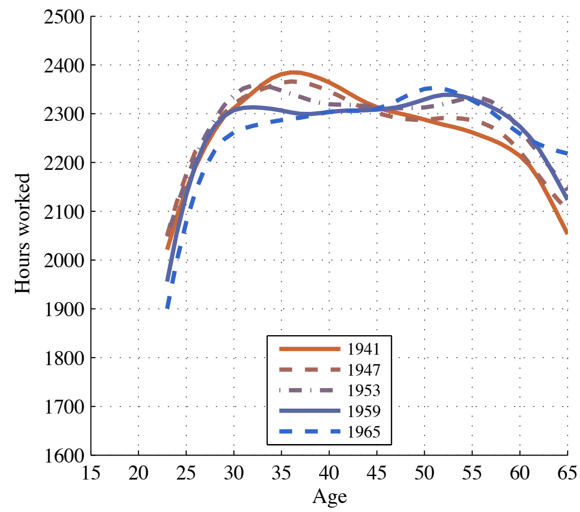
(a) High school dropouts



(b) High school graduates



(c) College dropouts



(d) College graduates

Figure 26: Age profiles of average annual hours worked for selected cohorts

Notes: CPS data.

A.4 Aggregate Statistics

Aggregate statistics for each year are computed for a population with constant age-school composition. First, I compute the average fraction of persons in each age-school cell across years. Then I create a constant-composition individual weight by multiplying the CPS person weights of all persons in a given age-school cell by a common factor such that the total weight of all persons in the cell adds up to the average weight of that cell across years.

Aggregate statistics, such as the standard deviation of log wages, are then computed by pooling all persons aged $25-55$ in a given year and applying the constant-composition weights.

B PSID Data

The data construction largely follows Hendricks (2013). The data are drawn from the $1968-1993$ waves of the Panel Study of Income Dynamics, distributed by the Institute for Social Research, Survey Research Center, University of Michigan (PSID). The sample includes male members of the core sample between the ages of 18 and 65 who report positive earnings and between 520 and 5110 work hours for at least 15 years.

Years of schooling are measured by the highest grade a person ever completed. Each person is assigned a school group based on their completed years of schooling using the following bounds: $<$12 years (HSD), 12 years (HSG), 13-15 years (CD), and 16+ years (CG). Earnings include wages and salaries as well as the labor part of business income. The wage is defined as earnings divided by annual work hours. Wages below $1/20$ or above 100 times the median wage in a given year are dropped as likely coding errors. Potential experience is set to $\text{age}-6-\max(12, \text{years of schooling})$.

The covariance matrix of log wages is constructed from the residual of a regression of log wages on a cubic polynomial in experience, pooling all cohorts. To increase sample sizes, the estimation pools high school dropouts with high school graduates and college dropouts with college graduates. It also groups 3 adjacent experience levels. Covariances are smoothed across leads using an HP filter. Estimating a first-order autoregressive process by minimizing the deviation between model and data covariances yields parameters that are close to those reported by Guvenen (2007), Table 1.

	All	HSD	HSG	CD	CG
N	5938	1523	2268	1159	988
Avg.school	13.0	9.8	12.0	13.8	16.9
Fraction	n/a	19.8	37.3	20.0	22.9
AFQT percentile		24.1	42.2	58.1	78.1
Mean experience	11.6	13.4	12.1	11.0	9.5
Max experience		39.0	30.0	29.0	26.0

Table 6: The NLSY79 sample

C NLSY79 Data

The data are drawn from the 1979-2006 waves of the National Longitudinal Survey of Youths 1979 cohort (Bureau of Labor Statistics; US Department of Labor, 2002). The sample contains men with positive sampling weight, known schooling and AFQT scores. Table 6 shows summary statistics.

Individual variables are constructed as follows. Years of schooling are set to the highest grade completed before the age of 30. School groups are coded from years of schooling using the same cutoffs as used for the PSID data. Wages are defined as labor earnings plus two-thirds of business income divided by hours worked per year. Wages are marked as invalid if hours worked fall outside the range 520-5110 or if they are below 1/20 or above 100 times the median wage in a given year. Potential experience is set to $\text{age} - 6 - \max(12, \text{years of schooling})$.