Chapter 1

POST SCHOOLING WAGE GROWTH: INVESTMENT, SEARCH AND LEARNING

YONA RUBINSTEIN

Brown University and Eitan Berglas School of Economics, Tel-Aviv University
e-mail: yona_rubinstein@brown.edu, yonar@post.tau.ac.il

YORAM WEISS

Eitan Berglas School of Economics, Tel-Aviv University
e-mail: weiss@post.tau.ac.il

Contents

Handbook of the Economics of Education, Volume 1
Edited by Eric A. Hanushek and Finis Welch

DOI: 10.1016/S1574-0692(06)01001-4

Abstract

The survey presents basic facts on wage growth and summarizes the main ideas on the possible sources of this growth. We document that wage growth happens mainly early in the life cycle and is then associated with increasing labor force participation and high job mobility. Wage growth during the first decade in the labor market, is about 50% for high school graduates and about 80% for those with college or more. This growth is comparable in size to the accumulated contribution of schooling for these two groups. We describe in detail models of wage growth that can explain these results, including investment in human capital, search and learning. We also discuss the roles of contracts in sharing the risks associated with learning about ability and varying market conditions. Evidence supporting investment is the U shaped life cycle profile for the variance of wages. However, heterogeneity matters and individuals with relatively high life time earnings have both a higher mean and a higher growth. Evidence supporting search is the high wage gains obtained from changing employers early in the career. Evidence for learning are the initially rising hazard of quitting and the rising rewards for AFQT scores that are not observed by the market.

1. Introduction

Perhaps the most widely estimated regression equation in economics is Mincer's log-earnings function that relates the log of individual earnings or wages to observed measures of schooling and potential work experience; with a specification that is linear in years of schooling and quadratic in experience. This simple regression has been estimated in numerous studies, employing various data sets from almost every historical period and country for which micro data are available, with remarkably robust regularities. First, workers' wage profiles are well ranked by education level; at any experience level, workers earn more, on average, as their schooling increases. Second, average wages grow at a decreasing rate until late in one's working lifetime. Most importantly, the estimated coefficients for schooling and experience in all these regressions fall into a sufficiently narrow range to admit a common economic interpretation in terms of rates of return for investment in human capital. The estimated coefficients of the log-earnings function have been applied to a wide variety of issues, including *ceteris paribus* effect of schooling on earnings, wage differentials by gender and race, and the evolution of earnings inequality. Mincer's (1974) earning function was used as the statistical platform in all these studies.[1]

The human capital approach to wage growth over the life cycle, as developed by Becker (1975), Mincer (1958, 1974) and Ben-Porath (1967), emphasizes the role of human capital acquired in school and on the job. Workers face a given trade off between current and future earnings, represented by a human capital "production function", and decide how much to invest. The wage offered to individuals is determined as a product of the worker's stock of human capital and the market-determined "rental rate". Markets operate competitively and workers are compensated for their investments. If individuals are heterogeneous, then compensation applies only at the margin, while non-marginal workers receive rents for their scarce attributes. When market conditions change, due to technological change for instance, the rental rate changes, as does perhaps the production function that describes the investment opportunities. Together, these lead to adjustments in the individual investment decisions that affect wage growth.

Becker (1975), Griliches (1977) and Rosen (1977) have questioned the interpretation that should be given to the regression coefficients of schooling and experience in the Mincer earning equation, and hence the validity of drawing policy conclusions from these coefficients. The main concerns are, first, the role of individual heterogeneity in ability and access to the capital markets and, second, the role of market frictions and specific investments in human capital. These concerns affect the statistical estimation procedures because the unobserved individual attributes that influence investment decisions can bias the schooling and experience coefficients in Mincer's equation. Equally important is the recognition that if markets are non-competitive because of credit constraints or the firm specific investments that create relational rents, then wages and

[1] Heckman, Lochner and Todd provide an insightful perspective on the Mincer earning regression, fifty years later.

productivity need not coincide as well as social and private rates of return for investment in human capital may diverge.

Parallel to the human capital approach, search models have been offered to deal with limited information and market frictions. At the individual level, these models explain wage growth and turnover as outcomes of the (random and intermittent) arrival of job offers that can be rejected or accepted [see Burdett (1978)]. These models also allow for investment in search effort, with the objective of generating job offers rather than enhancing productivity. When combined with learning, search models can provide a framework for explaining the separate roles of tenure and general market experience [see Mincer and Jovanovic (1981), Jovanovic (1984), Mortensen (1988)]. At the market level, search models can explain the aggregate level of unemployment in addition to the distribution of wages in the economy. The policy implications of these models for schooling and training may be quite different than those of the human capital model because of the important role of externalities, relational rents and bargaining [see Mortensen and Pissarides (1999), Wolpin (2003)].

A third important consideration that may explain wage growth is learning [see Jovanovic (1979a, 1979b), Harris and Holmstrom (1982), Gibbons and Waldman (1999a, 1999b)]. Workers are heterogeneous and it takes time to identify their productive capacity with sufficient precision. Therefore, employers must base their payments on predictions of expected output that are repeatedly modified by the worker performance. The arrival of new information which allows the market to sort workers can be individually costly, because it makes wages uncertain. This risk creates incentives for risk sharing between workers and firms. A possible outcome of this process is that all workers obtain partial insurance, to protect them against wage reductions upon failures to perform well. Yet, successful workers will be promoted because information is public and other firms compete for workers based on this information. We thus have wage growth that is triggered by new information rather than by the worker's actions or arrival of job offers.

Although investment, search and learning have similar implications with respect to the behavior of mean wages, implying rising and concave wage profiles, they can be distinguished by their different implications for higher moments, such as the wage variance. For instance, Mincer (1974) pointed out that compensation for past investment in human capital creates a negative correlation between early and late earnings during the life cycle, implying that the interpersonal variance of earnings over the life-cycle has a U-shape pattern. This is not true in the search and learning models, where workers that are initially homogeneous become increasingly heterogeneous as time passes due to their longer exposure to random job offers. In these models, the variance may first increase and then decrease as workers are gradually sorted into their "proper" place.

The purpose of this survey is to provide a synthesis of the alternative explanations for wage growth and relate them to the patterns observed in the data. The first part of the survey provides an initial glance at the data on life cycle wage levels and rates of wage growth, based on cross sectional, synthetic cohorts and panel data. We use all these sources to illustrate the important distinction between life-cycle and time effects and to show that most wage growth occurs early in the work career. These results are associated with high turnover, in and out of the labor force, between employers, occupations and industries. We show that post-schooling wage growth is quantitatively important and is as large as the wage growth attributed to schooling. Moreover, schooling and experience are strongly linked, with more-educated workers generally having higher wage growth and more-stable employment. The second part of the survey presents models of wage growth based on investment, search and learning in a unified framework. This allows us to compare alternative channels for wage growth and identify the connections amongst them. The third part of the survey provides a second glance, based on the empirical literature in the area and our own examination of the data, for the purpose of identifying empirical tests that take into account unobserved heterogeneity and might distinguish alternative models of wage growth.

2. Wages and employment over the life cycle – A first glance

In this section, we take a first glance at the available data on life cycle earnings. Our goal is to summarize the patterns of post schooling wages for workers of different educational attainments, without restricting ourselves to a particular functional form, such as the famous Mincer's wage equation that restricts mean (log) wages to be linear in schooling and quadratic in experience. We take advantage of large bodies of data collected over several decades, a privilege that early research did not have, for reproducing the basic facts on wages over the life cycle.

The data sources are the March Supplements from the Current Population Surveys (CPS) for the years 1964–2002, the Panel Study of Income Dynamics (PSID) for the years 1968–1997 the National Longitudinal Survey of Youth (NLSY) for the years 1979–2000, and the CPS outgoing rotation groups (ORG) for the years 1998–2002.

The March CPS data is a sequence of annual cross sections. The ORG CPS data follows households over 16 months and enables us to create short panels for individuals. The PSID began with a cross-sectional national sample in 1968, with participants interviewed every year until 1993 and then biannually until 1997. In contrast, the NLSY sample includes only individuals aged 14–21 when first interviewed in 1979 and observed until 2000. (A more detailed description of these data sets is available in the Appendix.)

From each source, we selected white males with potential work experience (age – school years – 6) of no more than 40 years. Observations were divided by school completion into five levels: (i) high school dropouts, (ii) high school graduates with twelve years of schooling, (iii) some college, (iv) college graduates with a BA degree and (v) college graduates with advanced/professional education (MBA, PhD). We then examine the hourly or annual wages, whichever is applicable, of workers employed full time and full year.

By restricting ourselves to white US males, we can examine wage patterns for a relatively homogeneous group over a long period of time. This allows us to control for institutional and social differences and to focus on the potential role of the economic forces that affect wage growth, such as investment, search and prices of skills.

2.1. The pooled data

Under stationary conditions, the chronological time of observation would be irrelevant; we can then pool data from different years and cohorts while paying attention only to the stage in the worker's life cycle, as indicated by his potential work experience. Figure 1 shows the mean weekly wage–experience profiles, by schooling, averaged over the 38 years 1964 to 2002 of the March CPS data, using a subsample of fully employed (full time and full year) workers. These (log) wage profiles have the general shape found in previous studies based on single cross sections [see Mincer (1974), Murphy and Welch (1992), Heckman, Lochner and Todd (2001)]. Average wages are well ranked by educational attainments. Mean wages increase rapidly (by approximately 80 percent) over the first 10 to 15 years of a career. As careers progress, we find little change in mean wages.

The sharp growth in wages is associated with a sharp increase in labor supply and regularity of employment, as indicated by the life-cycle profiles of the proportion of workers who work full time, full year (among those who worked some time during the year) and average weekly hours (for those with positive hours). Workers with higher levels of schooling work more and reach a steady level much earlier than do less educated workers (see Figures 2a and 2b). Thus, hours and wages move together over the life cycle, and earnings grow faster than wages.

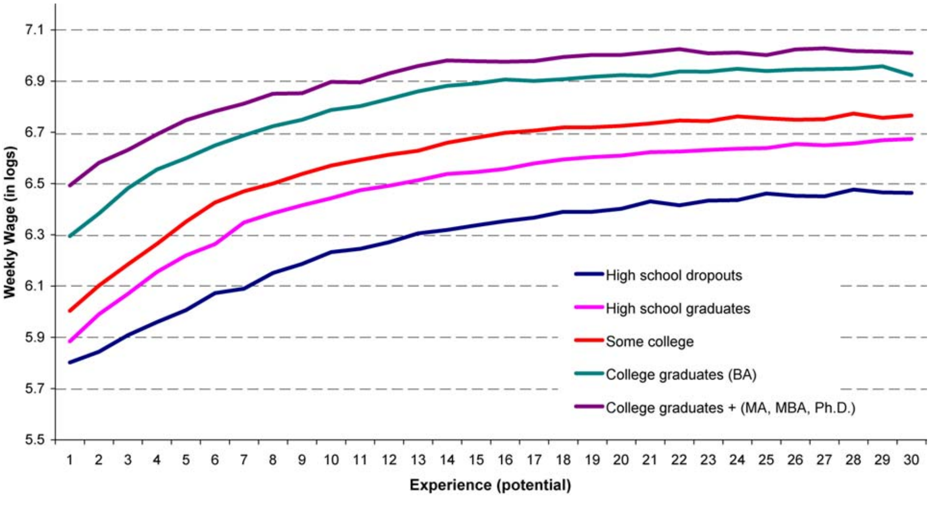

Figure 1. Mean weekly wages (in logs) by education and (potential) experience, white males, full-time full-year workers (52 weeks), CPS, March supplement, 1964–2002.

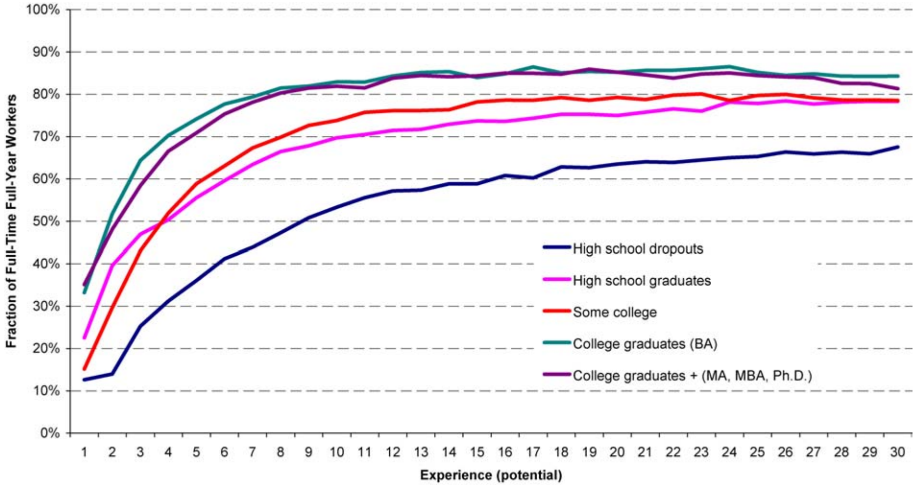

Figure 2a. Fraction of full-time full-year workers and average weekly hours of employed workers by education and experience, CPS, March supplement, 1964–2002. Fraction of full-time full-year workers.

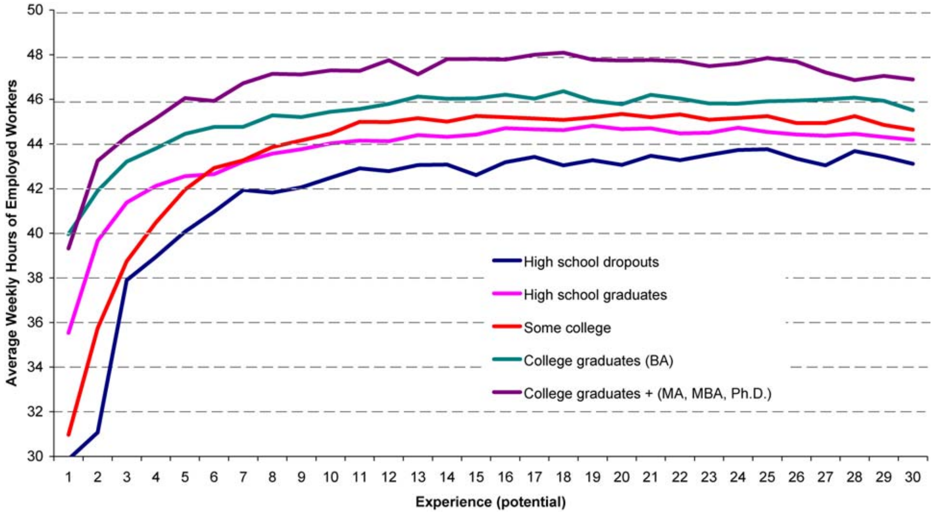

Figure 2b. Fraction of full-time full-year workers and average weekly hours of employed workers by education and experience, CPS, March supplement, 1964–2002. Average weekly hours of employed workers.

2.2. Cohorts and cross-sections

In fact, the economy is not stationary. The wage structure has undergone major changes beginning in the late 1970's, when workers with high level of schooling started to gain relative to those with low levels of schooling, mainly as a result of the decline in the wages of low-skill workers [see Katz and Autor (1999)]. Such changes in returns to skill imply different wage profiles for different cohorts, where workers born in the same year are followed over time, and for cross sections, where workers with different experience (and time of entry into the labor force) are observed at a given year.

Figures 3a and 3b show the wage–experience profiles for the cohort of high school graduates born in 1951–1955 and the cohort of college graduates born in 1946–1950, respectively. These two groups entered the labor market at roughly the same time, 1971–1975. Added to the graphs is the evolution of the cross section wage–experience profiles from 1971 to 2000 in five year intervals, where each such cross section profile shows the mean wages of workers with the indicated schooling and experience in a given time interval. These figures make it very clear that cohort-based wage profiles are affected by changes in market conditions that shift the cross section profiles over time. These shifts differ by level of schooling. High school graduates of *all* experience levels earned lower wages during the period 1970–2000, which is the reason why the mean wage profile of the cohort of high school graduates born between 1951 to 1955 exhibits almost no wage growth after ten years in the labor market (see Figure 3a). In contrast, workers with a college degree or more maintained their earning capacity over time. Consequently, as seen in Figure 3b, the cross section and cohort wage profiles of college graduates are quite similar and rise throughout most of the worker's career.

Although the cross section profile is, by construction, free of time effects, its shape is not necessarily a reflection of life cycle forces because cohorts "quality" can change over time. An important reason for this is that schooling is embodied in the worker early in life and the quality of that schooling may depend on the size of the cohorts with each level of schooling and the state of knowledge at the time of entry. It is impossible to separately identify time cohort and life cycle effects unless one uses some a priori identifying assumptions.[2]

2.3. Panel data

Panel data follows the *same* group of individuals over a period of time, in contrast to cohort data, where different individuals are sampled in every period. Having repeated observations for the same individual allows one to calculate individual rates of wage growth and examine their variance. The panel also allows examination of individual transitions among different employers and occupations.

[2] For instance, Borjas (1985) assumed that time effects are common to immigrants and natives to identify cohort effects for immigrants. Weiss and Lillard (1978) assumed that time effects are constant and common to all experience groups in order to identify cohort effects for scientists.

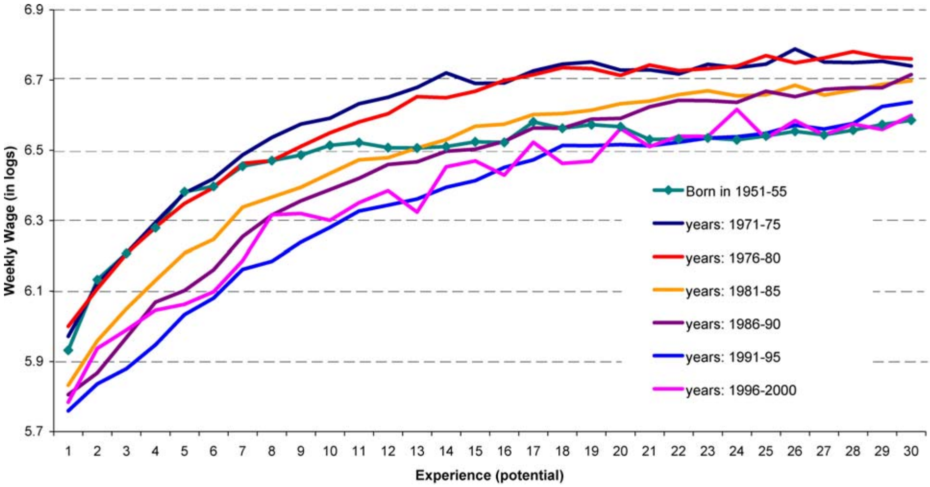

Figure 3a. Cohort and cross-section wage profiles for high school graduates and college graduates, white males, CPS, March supplement, 1964–2002. High school graduates.

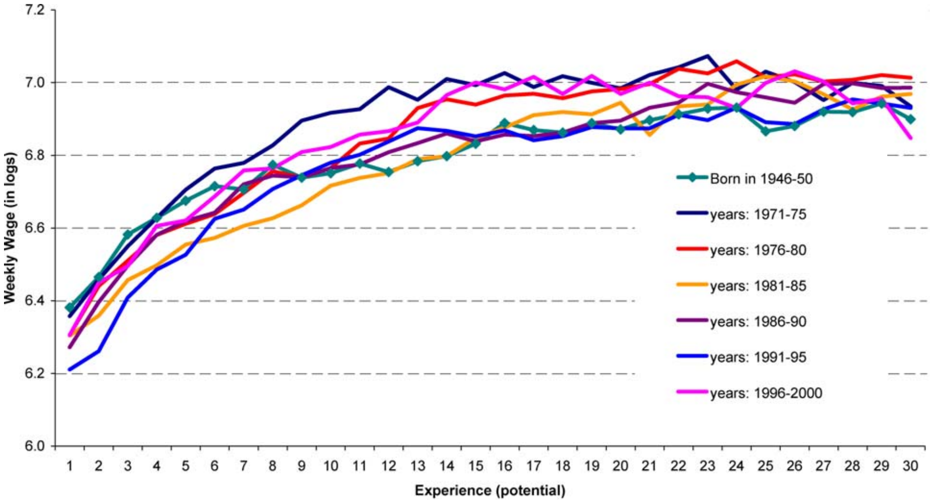

Figure 3b. Cohort and cross-section wage profiles for high school graduates and college graduates, white males, CPS, March supplement, 1964–2002. College graduates.

Figures 4a and 4b show the average wage profiles constructed from PSID and NLSY data. Basically, the patterns resemble the synthetic cohorts displayed in Figures 3a and 3b, except that the panel profiles are less likely to taper off and decline late in the life cycle for workers with less than a college degree. Note that the NLSY sample

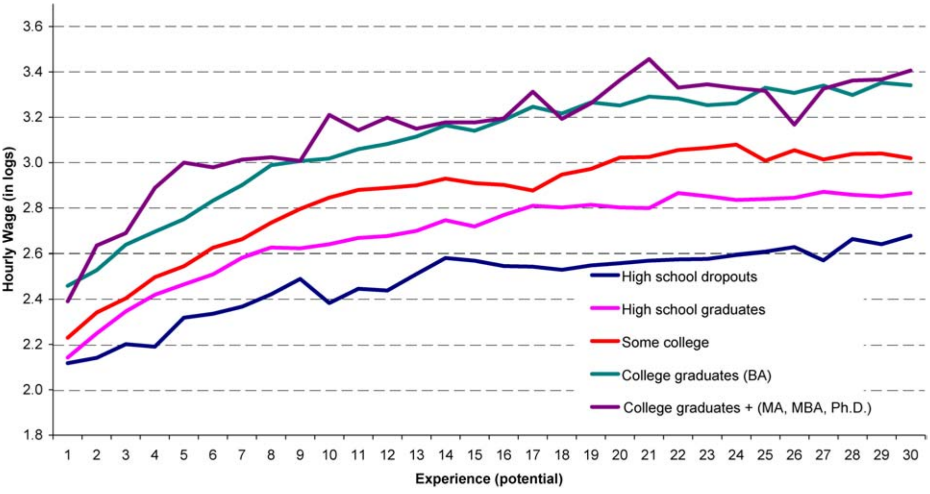

Figure 4a. Mean hourly wages (in logs) by education and experience, PSID, 1968–1997 and NLSY, 1979–2000. PSID, 1968–1997.

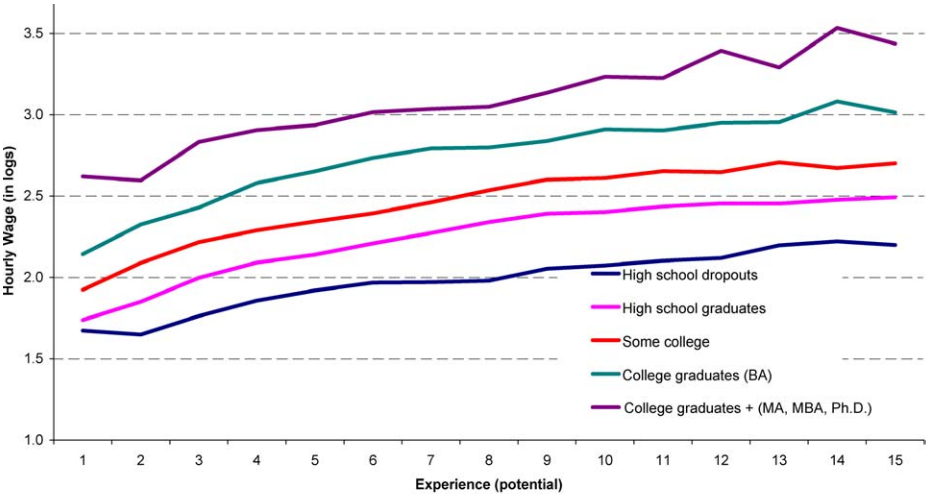

Figure 4b. Mean hourly wages (in logs) by education and experience, PSID, 1968–1997 and NLSY, 1979–2000. NLSY, 1979–2000.

follows few birth cohorts that are close to each other, at the early stage of the life-cycle, while the PSID covers many cohorts at all stages of the life cycle. Therefore, the NLSY profiles are less concave than the corresponding PSID profiles, which show a pattern that is more similar to the CPS cross section profiles.

Figures 5a and 5b display the life cycle patterns of the monthly proportions of CPS workers that changed occupation and industry, while Figure 5c shows the annual proportions of NLSY workers who changed employers. We see that for all these dimensions of mobility, transitions decline quickly with potential experience and are generally more frequent among the less educated, especially at the early part of their careers. The impact of schooling on movement across employers is weaker than on transitions across occupations or industries. Similar findings are reported by Topel and Ward (1992), Hall (1982), Blau and Kahn (1981), Mincer and Jovanovic (1981), Abraham and Farber (1987), Wolpin (1992) and Farber (1999).

An interesting feature of the transitions among employers is that the proportion of movers initially rises, suggesting a period of experimentation on the job, and continues at a relatively high rate of about 15 percent per year until the end of the worker's career.

2.4. Individual growth rates

Table 1 summarizes the main results on wage growth. For each individual, we calculate annual wage growth and then present the averages and standard deviations of these rates, by experience and schooling. For comparison, we also present the predicted average growth rates that would be implied for the same individuals by using Mincer's quadratic specification for wage levels. We report these figures for the CPS short panel as well as the PSID and the NLSY samples. We include only observations in which workers were fully employed in the two consecutive years for which wage growth is calculated (see Appendix).

The average worker's career is characterized by three very different phases. The first, decade-long phase is characterized by a sharp growth of wages. The second, five-year long phase is characterized by moderate wage growth; the late phase of a career has zero or negative growth. The growth rates are substantially higher for workers with higher levels of schooling. This general pattern is revealed in *all* the data sets that we use. However, the CPS short panel shows somewhat lower rates of wage growth because of the absence of time effects.

The average annual growth rates of wages in the initial ten years for the most-educated group are 7.7 in the CPS short panel, and 11.0 and 9.6 in the PSID and NLSY panels, respectively. These rates are quite close to the wage growth associated with schooling. However, the contribution of experience declines with the level of schooling; for high school graduates, average growth rates during the first decade of post schooling experience are 5.6, 5.7 and 7.1 in the CPS, PSID and NLSY, respectively.

There is a sharp decrease in wage growth with labor market experience. As one moves across experience groups for the highly educated, the wage growth in the CPS short panel declines from 7.7 to 5.3 and then to 1.5. In the PSID sample, wage growth declines from 11.0 to 1.3 and then rises stightly to 1.9. The NLSY sample shows no such reduction mainly because it represents few cohorts, all of which gain from the continuous rise in skill prices. For some college and below, we see a decline of wage growth with experience in all samples because these groups gained less from the increase in skill prices.

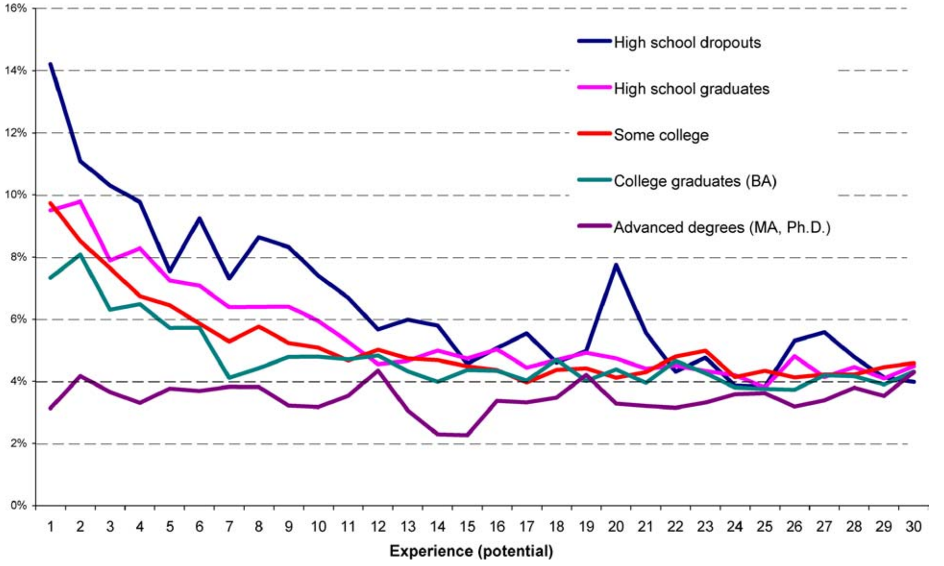

Figure 5a. Proportion of workers who changed occupation, industry or employers by education and experience, full-time workers, CPS-ORG, 1998–2002, and NLSY, 1979–2000. Proportion of workers who changed occupation (within one month), CPS-ORG, 1998–2002.

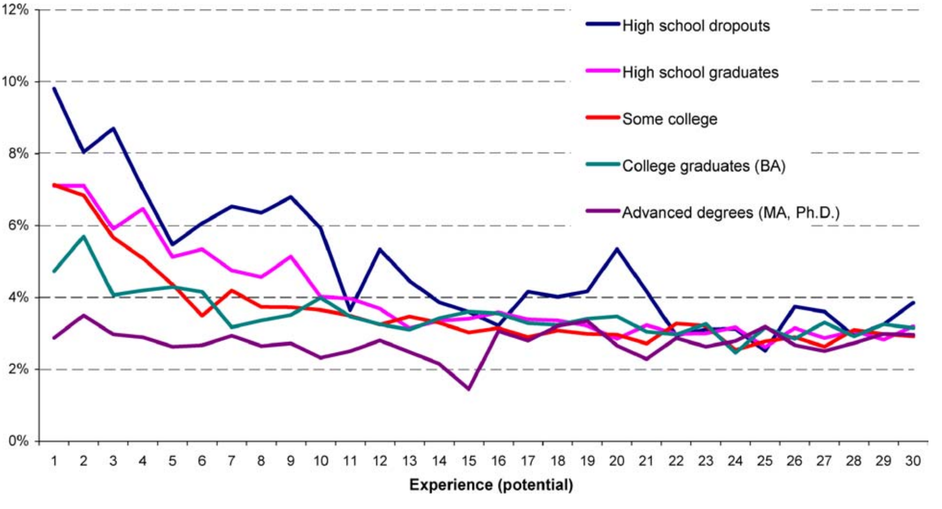

Figure 5b. Proportion of workers who changed occupation, industry or employers by education and experience, full-time workers, CPS-ORG, 1998–2002, and NLSY, 1979–2000. Proportion of workers who changed industry (within one month), CPS-ORG, 1998–2002.

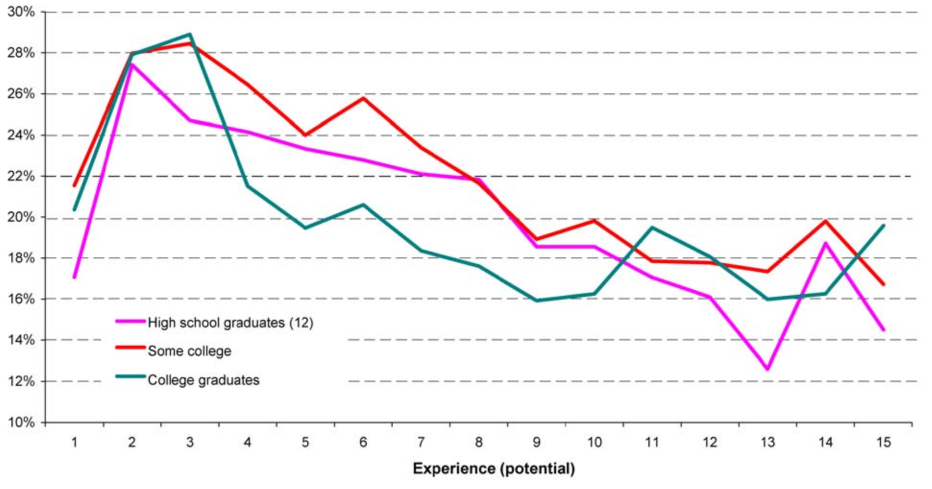

Figure 5c. Proportion of workers who changed occupation, industry or employers by education and experience, full-time workers, CPS-ORG, 1998–2002, and NLSY, 1979–2000. Proportion of workers who changed employers (within one year), NLSY, 1979–2000.

Differences in average growth rates by schooling levels are substantial. For instance, in the CPS and PSID samples, workers with advanced degrees enjoy a wage growth that is twice as high as that of workers with less than high school degree (.077 vs. .039 and .110 vs. .043, respectively) during the first decade of their career. This important interaction is not captured by the standard Mincer specification; we allow for it here because we estimate the experience coefficients separately for each education group. As seen in Table 1, the averaged individual growth rates are generally higher than the wage growth obtained from Mincer's quadratic specification, especially at the early part of a career. As noted by Murphy and Welch (1990), the quadratic specification overestimates early wages and underestimates late wages. As a consequence of this misspecification, early growth rates are substantially biased downwards.

The variability in the rates of wage growth follows a U-shape pattern with respect to schooling. That is, the standard deviations are lower for workers with high school degree than for workers with more schooling or less, suggesting that, in this regard, the middle levels of schooling are less risky. However, there is no systematic pattern for the standard deviations of wage growth by level of experience.

In Table 2a we show, for each experience and education group, the proportion of observations with a rise, a decline and no change in *reported nominal* wage;[3] for each such subsample, we calculate the average change in *real* hourly wage. Using the CPS short panel, we see that, given a nominal increase, the average real hourly wage grows at a

[3] The wage used for this classification is total annual salary reported in the NLSY and PSID. For the CPS short panel, we use the monthly wage. These are raw data and no correction for hours was made.

Table 1
The average wage growth by education, experience, specification and data sources

Experience	Data source	Education categories									
		Less than HSG		HSG (12)		Some college		College graduates		MA, Ph.D.	
		Level	Dif	Level	Dif	Level	Dif	Level	Dif	Level	Dif
0–10											
	CPS-ORG	0.024	0.039	0.032	0.056	0.033	0.063	0.036	0.063	0.029	0.077
		(0.003)	(0.029)	(0.001)	(0.010)	(0.001)	(0.010)	(0.002)	(0.011)	(0.003)	(0.017)
	PSID	0.028	0.043	0.030	0.057	0.038	0.065	0.039	0.076	0.032	0.110
		(0.003)	(0.007)	(0.002)	(0.003)	(0.003)	(0.005)	(0.003)	(0.004)	(0.006)	(0.021)
	NLSY	0.024	0.065	0.034	0.071	0.046	0.081	0.052	0.082	0.055	0.096
		(0.006)	(0.010)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.009)	(0.012)
11–15											
	CPS-ORG	0.016	0.007	0.022	0.033	0.022	0.055	0.022	0.045	0.018	0.053
		(0.002)	(0.034)	(0.001)	(0.011)	(0.001)	(0.012)	(0.001)	(0.012)	(0.001)	(0.020)
	PSID	0.019	0.030	0.020	0.021	0.026	0.021	0.027	0.029	0.022	0.013
		(0.002)	(0.007)	(0.001)	(0.004)	(0.002)	(0.005)	(0.002)	(0.005)	(0.004)	(0.016)
	NLSY	0.013	0.024	0.023	0.019	0.026	0.024	0.035	0.067	0.039	0.123
		(0.002)	(0.008)	(0.001)	(0.004)	(0.002)	(0.007)	(0.004)	(0.009)	(0.009)	(0.018)
16–25											
	CPS-ORG	0.010	0.052	0.013	0.022	0.012	0.026	0.009	0.026	0.009	0.015
		(0.001)	(0.021)	(0.000)	(0.007)	(0.000)	(0.008)	(0.001)	(0.009)	(0.001)	(0.012)
	PSID	0.011	0.010	0.012	0.010	0.015	0.014	0.017	0.026	0.014	0.019
		(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)	(0.004)	(0.003)	(0.009)
	NLSY	0.003	0.035	0.014	0.038	0.009	0.065	0.021	0.111	0.025	0.044
		(0.004)	(0.009)	(0.003)	(0.005)	(0.005)	(0.013)	(0.009)	(0.015)	(0.022)	(0.035)
25+											
	CPS-ORG	−0.002	0.025	−0.004	0.011	−0.005	0.002	−0.014	−0.002	−0.009	0.012
		(0.003)	(0.017)	(0.001)	(0.007)	(0.001)	(0.008)	(0.002)	(0.011)	(0.003)	(0.013)
	PSID	−0.003	0.004	−0.005	0.006	−0.005	0.010	−0.003	0.000	−0.001	0.011
		(0.001)	(0.003)	(0.002)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
	NLSY	−0.015	0.034	−0.003	0.034						
		(0.012)	(0.042)	(0.007)	(0.041)						

Notes: The numbers in the "dif" columns are cell means and standard deviations. The numbers in the "level" columns are growth rates as implied by the estimated coefficients of the experience and experience squared terms in Mincer's wage equation.

hefty rate of 25 percent per year. The corresponding figure for wage reduction is even larger, −33 percent per year. As experience increases, the proportion of gainers (workers with a wage rise) declines and the proportion of losers (workers with a wage decline) rises. However, the conditional means of their respective wage changes remain remarkably similar across experience groups. Similarly, as we compare education groups, the main reason for the higher growth rate among the educated is the larger proportion of

Table 2a
Annual wage growth rates and proportions of gainers and losers, by education and experience; CPS-ORG, 1998–2002

		High school graduates		Some college		College graduates		Advanced degrees	
	Experience	Fraction	Wage growth	Fraction	Wage growth	Fraction	Wage growth	Fraction	Wage growth
All	0–10	1.000	0.056	1.000	0.063	1.000	0.063	1.000	0.077
	11–15	1.000	0.033	1.000	0.055	1.000	0.045	1.000	0.053
	16–25	1.000	0.022	1.000	0.026	1.000	0.026	1.000	0.015
	26–40	1.000	0.011	1.000	0.002	1.000	−0.002	1.000	0.012
Gainers (wage up)	0–10	0.602	0.259	0.621	0.255	0.643	0.263	0.667	0.253
	11–15	0.588	0.254	0.589	0.254	0.602	0.259	0.590	0.274
	16–25	0.562	0.264	0.582	0.257	0.567	0.268	0.567	0.250
	26–40	0.546	0.264	0.555	0.261	0.536	0.287	0.545	0.265
No wage change	0–10	0.048	−0.022	0.043	−0.035	0.055	−0.025	0.080	−0.003
	11–15	0.048	−0.017	0.056	0.001	0.081	−0.026	0.090	−0.038
	16–25	0.049	−0.018	0.053	−0.036	0.090	−0.031	0.083	−0.007
	26–40	0.053	−0.028	0.055	−0.037	0.085	−0.026	0.099	−0.020
Losers (wage down)	0–10	0.349	−0.289	0.336	−0.279	0.301	−0.350	0.253	−0.361
	11–15	0.363	−0.306	0.355	−0.267	0.317	−0.343	0.320	−0.329
	16–25	0.389	−0.325	0.364	−0.334	0.342	−0.359	0.349	−0.363
	26–40	0.401	−0.314	0.391	−0.361	0.378	−0.405	0.356	−0.366

Notes: Gainers (losers) had a nominal wage increase (decrease) between subsequent wage observations.
Fraction is the share within experience groups.

Table 2b
Annual wage growth and proportions of gainers and losers by experience groups and data source

	Experience	CPS-ORG Fraction	CPS-ORG Wage growth	NLSY Fraction	NLSY Wage growth	PSID Fraction	PSID Wage growth
All							
	0–10	1.000	0.062	1.000	0.077	1.000	0.063
	11–15	1.000	0.044	1.000	0.033	1.000	0.024
	16–25	1.000	0.024	1.000	0.049	1.000	0.015
	26–40	1.000	0.007	–	–	–	–
Gainers (wage up)	0–10	0.627	0.259	0.718	0.176	0.726	0.163
	11–15	0.593	0.254	0.644	0.144	0.689	0.122
	16–25	0.568	0.264	0.662	0.168	0.667	0.118
	26–40	0.547	0.264	–	–	–	–
No wage change	0–10	0.053	−0.023	0.071	−0.044	0.040	−0.040
	11–15	0.065	−0.020	0.097	−0.040	0.048	−0.030
	16–25	0.065	−0.023	0.082	−0.041	0.056	−0.046
	26–40	0.066	−0.029	–	–	–	–
Losers (wage down)	0–10	0.319	−0.312	0.211	−0.221	0.234	−0.228
	11–15	0.342	−0.309	0.259	−0.217	0.263	−0.224
	16–25	0.367	−0.339	0.255	−0.232	0.277	−0.220
	26–40	0.388	−0.351	–	–	–	–

Notes: Gainers (losers) had a nominal wage increase (decrease) between subsequent wage observations. Fraction is the share within experience groups.

workers with a nominal wage rise; but given such a change, the average increase is *independent* of the level of schooling.

The same patterns are seen in Table 2b for the NLSY and PSID samples, where due to the smaller size of these samples we classify the data only by experience. Again, the main reason for the reduction of wage growth with experience is the decline in the proportion of gainers, while the conditional means remain the same (except for gainers in the PSID who show some decline).

Finally, Table 2c shows the interaction between gainers, losers, movers and stayers. It is seen that, compared to stayers, workers who change employers are more likely to be losers and suffer a larger reduction in wages if they lose. However, movers obtain higher wage increases if they gain. In this respect, the current job provides workers with some insurance. Taken together, the patterns displayed in Figure 3 strongly suggest that the average wage growth is influenced by the arrival of positive or negative shocks. It is the nature of such shocks (positive or negative) rather than their size that changes over the life cycle.

Table 2c
Annual wage growth and proportions of gainers, losers, movers and stayers in the NLSY, by experience groups

	Experience	All		Stayers		Movers	
		Fraction	Wage growth	Fraction	Wage growth	Fraction	Wage growth
All	0–10	1.000	0.077	0.800	0.082	0.200	0.052
	11–15	1.000	0.033	0.833	0.039	0.167	0.001
	16–25	1.000	0.049	0.833	0.055	0.167	0.014
Gainers (wage up)	0–10	0.718	0.176	0.739	0.170	0.625	0.208
	11–15	0.644	0.144	0.662	0.140	0.549	0.174
	16–25	0.662	0.168	0.680	0.162	0.568	0.207
No wage change	0–10	0.071	−0.044	0.070	−0.046	0.074	−0.034
	11–15	0.097	−0.040	0.100	−0.038	0.080	−0.049
	16–25	0.082	−0.041	0.083	−0.040	0.079	−0.045
Losers (wage down)	0–10	0.211	−0.221	0.191	−0.210	0.301	−0.250
	11–15	0.259	−0.217	0.238	−0.209	0.370	−0.244
	16–25	0.255	−0.232	0.237	−0.218	0.353	−0.283

Notes: Gainers (losers) had a nominal wage increase (decrease) between subsequent wage observations.
Movers (stayers) changed (did not change) employer between subsequent wage observations.
Fraction is the share within experience groups.

2.5. *The questions*

Based on this preliminary glance at the data, the following questions arise:

- What causes the large wage growth at the initial phase of a career?
- Why does wage growth decline?
- What are the interrelationships between wage growth, job change and labor supply?
- What causes the large variance in individual wage growth and who are the gainers and losers?

In the next section, we examine some theoretical models that address these issues. In the subsequent (and last) section, we present further evidence and discuss the support for these explanations that is provided by the data.

3. Models of wage growth

A basic tenet of modern labor economics is that the observed life cycle wage patterns are, to a large extent, a matter of choice. Thus, each worker can influence his future wage by going to school, by choosing an occupation and by searching for a better job. Of course, wage levels and wage growth are also influenced by factors beyond the worker's control, such as aggregate demand and supply, technology, degree of competition and the institutional framework. Nevertheless, individual choice in a given market situation is an important part of the equilibrium analysis of wage outcomes.

In this survey, we present some of the basic approaches that economists have used in the analysis of post-schooling wage growth. The main ideas that we cover are investment search, and learning. Our purpose is to illustrate how these ideas are used in sufficient detail to enable the reader to use them as tools. We try to use as much a unified framework, as possible, so as to make the conceptual connections and distinctions between these ideas transparent. To achieve this purpose within our space constraints we have omitted important ideas that require separate discussion. In particular, we focus on general training and do not discuss firm-specific investments, mainly because of the difficulties in pinning down the wages. We also do not cover incentive contracts and the relations between wages and effort. The interested reader should consult other surveys for these important and complex issues [Malcomson (1997, 1999), Gibbons and Waldman (1999a, 1999b), Prendergast (1999)]. Finally, we do not discuss the important relationships between wages and hours worked [see Weiss (1986), Blundell and MaCurdy (1999)].

3.1. Investment

Workers have a finite life, T, and time is discrete. Let Y_t denote the earning capacity of the worker with the *current* employer, t, $t = 1, 2, \dots, T$. We assume that

$$Y_t = R_t K_t, \tag{1}$$

where K_t is the worker's human capital and R_t is the rental rate. In a competitive world, without friction, all firms pay the *same* rental rate.

Workers can accumulate human capital by investment on the job. Let l_t be the proportion of earnings capacity that is forgone when the worker learns on the job. Hence, current earnings are

$$y_t = R_t K_t (1 - l_t). \tag{2}$$

Following the Ben-Porath (1967) model, suppose that human capital evolves according to

$$K_{t+1} = K_t + g(l_t K_t), \tag{3}$$

where $g(\cdot)$ is increasing and concave with $g(0) = 0$. Thus, a worker who directs a larger share of his existing capital to investment has lower current earnings but a higher future earning capacity.

Here we consider only the behavior of workers for a given "production function" $g(\cdot)$. In a more general analysis, this function would be influenced by market forces [see Rosen (1972), Heckman, Lochner and Taber (1998)], but we do not attempt to close the model by deriving the equilibrium trade-off between current and future earnings.

To determine a worker's investment, we form the Bellman equation

$$V_t(K_t) = \operatorname*{Max}_{l_t}\left[R_t K_t(1 - l_t) + \beta V_{t+1}\left(K_t + g(l_t K_t)\right)\right], \tag{4}$$

where β represents the discount factor and $\beta < 1$. This equation states that the value of being employed in period t consists of the current earnings with this employer and the *option* to augment human capital through learning on the job. Each of these terms depends on the level of investment of the worker, and one considers only the optimal choices of the worker in calculating the value of the optimal program.

The first-order condition for l_t in an interior solution is

$$\frac{R_t}{g'(l_t K_t)} = \beta V'_{t+1}(K_{t+1}). \tag{5}$$

The left-hand side of (5) describes the marginal costs of investment in terms of forgone current earnings, while the right-hand side is the marginal value of additional future earnings. In the last period, T, investment is zero because there are no future periods left in which to reap the benefits.

Differentiating both sides of (4) w.r.t. K_t and using (5) we obtain the rule of motion for the marginal value of human capital

$$V'_t(K_t) = R_t + \beta V'_{t+1}(K_{t+1}). \tag{6}$$

Using the end condition that $V_{T+1}(K_{T+1}) = 0$ for *all* K_{T+1}, meaning that human capital has no value beyond the end of the working period, we obtain

$$V'_T(K_T) = R_T. \tag{7}$$

The standard investment model assumes stationary conditions; hence, R_t is a constant that can be normalized to 1. Then, using (7) and solving (6) recursively, the value of an additional unit of human capital at time t is

$$V'_t(K_t) = \frac{1 - \beta^{T+1-t}}{1 - \beta}, \tag{8}$$

which is *independent* of K_t. It follows that the value of being employed at a given current wage *declines* with time, that is, $V'_t(K_t) \geqslant V'_{t+1}(K_{t+1})$ for all periods $t = 1, 2, \ldots, T$. The shorter the remaining work horizon, the less valuable is the current stock of human capital and the lower the incentive to augment that stock. The lack of dependence on history, implicit in the Ben-Porath (1967) specification, is sufficient but not necessary for the result of declining investment, which holds under more general conditions [see Weiss (1986)].

The model can be easily generalized to the case in which R_t is variable over time. In this case, equation (8) becomes

$$V'_t(K_t) = \sum_{\tau=t}^{T} \beta^{\tau-t} R_\tau. \tag{8$'$}$$

Comparing these expressions, it is seen that if R_t rises with time, then the investment in human capital is higher at each period. The reason is that investment occurs when a worker receives a relatively lower price for his human capital, so that the forgone earnings are relatively low. If the rental rate rises with time at a decreasing rate, this relative price effect weakens with time and investment declines.[4]

The observable implications of this model are clear:

- For a constant R, investment declines as the worker ages and approaches the end of his working life.
- Earnings rise along an optimal investment path. This is caused by two effects that reinforce each other; positive investment increases earning capacity and declining investment induces a rise in its utilization rate.
- If R varies with time, workers that expect exogenous growth in their earning capacity invest at a higher rate and their wage rises at a higher pace. Investment declines if the rate of growth in the rental rate decreases.

3.2. *Investment in school and on the job*

Investment in school and on the job can be viewed as two alternative modes of accumulation of human capital that complement and substitute each other. Complementarity arises because human capital is self-productive, so that human capital accumulated in school is useful for learning on the job. Substitution arises because life is finite and if more time is spent in school, there is less time left for investment on the job. Although the focus of this survey is on post-schooling investments, the fact that these two modes are to some extent jointly determined leads us to expect interactions, whereby individuals completing different levels of schooling will invest differentially on the job and therefore display different patterns of wage growth.

Investment on the job is usually done *jointly* with work, while schooling is done separately. As a consequence, one foregoes less earning when training on the job than in school. However, in school, one typically specializes in the acquisition of knowledge and human capital is consequently accumulated at a faster rate. One can capture these differences by assuming different production (and cost) functions for the two alternative investment channels.

Let p_t be a labor force participation indicator such that $p_t = 1$ if the individual works in period t and $p_t = 0$, otherwise. Suppose that when the individual does not work he

[4] Using (8′) and (5) investment is determined by

$$\frac{1}{g'(l_t K_t)} = \sum_{\tau=t}^{T} \beta^{\tau-t} \frac{R_\tau}{R_t}.$$

If the rental rate R rises with time $\frac{R_\tau}{R_t} > 1$, which raises the incentives to invest at any period. If, in addition, $\frac{R_\tau}{R_t}$ declines in t for all $\tau > t$ then changing prices creates an added incentive to invest early rather than late, which together with the effect of the shortening horizon implies that investment, $l_t K_t$, declines in t.

goes to school and then accumulates human capital according to $K_{t+1} = K_t(1+\gamma)$ where γ is a fixed parameter such that $\gamma K_t > g(l_t K_t)$. We also assume that $(1+\gamma) > \frac{1}{\beta}$, which means that the rate of return from investment in human capital γ exceeds the interest rate. Otherwise, such investment would never be optimal. Assume stationary conditions and let $R_t = 1$. We can now rewrite the Bellman equation in the form

$$V_t(K_t) = \underset{p_t, l_t}{\mathrm{Max}}\left[p_t K_t(1-l_t) + \beta V_{t+1}\big(K_t + p_t g(l_t K_t) + (1-p_t)\gamma K_t\big)\right]. \tag{9}$$

School is the preferred choice in period t if

$$\beta V_{t+1}\big(K_t(1+\gamma)\big) > K_t(1-l_t^*) + \beta V_{t+1}\big(K_t + g(l_t^* K_t)\big), \tag{10}$$

where the optimal level of training on the job, l_t^*, is determined from (5). Finally, the law of motion for the marginal value of human capital is modified to

$$V_t'(K_t) = p_t + \beta V_{t+1}'(K_{t+1})\big(1 + (1-p_t)\gamma\big). \tag{11}$$

This extension has several implications:

- Specialization in schooling occurs, if at all, in the first phase of life. It is followed by a period of investment on the job. In the last phase of the life cycle, there is no investment at all.
- During the schooling period, there are no earnings, yet human capital is accumulated at the maximal rate $(1+\gamma)$. During the period of investment on the job, earnings are positive and growing. In the last phase (if it exists), earnings are constant.
- A worker leaves school at the first period in which (10) is reversed. At this point it must be the case that $l_t^* < 1$, which means that at the time of leaving school, earnings must *jump* to a positive level. This realistic feature is present only because we assume different production (and cost) functions in school and on the job, whereby accumulation in school is faster but requires a larger sacrifice of current earnings.
- A person with a larger initial stock of human capital, K_0, will stay in school for a shorter period and spend more time investing on the job. He will have higher earnings and the same earnings growth throughout life.
- A person with a larger scholastic learning ability, γ, will stay in school for a longer period and spend less time investing on the job. He will also have higher earnings and the same earning growth throughout life.

Although these results depend heavily on the particular form of the production function (3), they illustrate that unobserved characteristics of economic agents can create a negative correlation between the amounts of time spent investing in school and on the job, while there need be no correlation between completed schooling and post schooling wage growth.[5] It should be noted, however, that wage growth is often higher for the more educated, which casts some doubt on the neutrality implied by (3). Uncertainty and unexpected shocks can also affect the correlation between schooling and investment. For instance, the introduction of computers may raise the incentive to invest on the job among educated workers to a larger extent than among uneducated workers because the investment's payoff may be lower for the second group.[6]

3.3. Search

In a world with limited information and frictions, firms may pay a different rental rate, R, because workers cannot immediately find the highest paying firm and must spend time and money to locate employers. If a worker meets a new employer, he obtains a random draw $\widetilde{R}$ from the given distribution of potential wage offers $F(R)$. The worker decides whether to accept or reject this offer. To simplify, we assume here that workers are relatively passive in their search for jobs. They receive offers at some fixed exogenous rate λ, but do not initiate offers through active job search.

We discuss here the case with homogenous workers and firms, assuming that workers are equally productive in all firms and their productivity is constant over time. However, firms may pay different wages for identical workers. Specifically, if K is the worker's human capital, then the profits of a firm that pays the worker R are $K - RK$. Firms that post a high R draw more workers and can coexist with a firm that posts a low R and draws few workers. In equilibrium, all firms must have the same profits [see Mortensen and Pissarides (1999)]. Here we consider only the behavior of workers for a given wage distribution, $F(R)$, and do not attempt to close the model by deriving either the equilibrium wage offer distribution or the equilibrium trade-off between current and future earnings. In a more general analysis, the wage distribution is determined by market forces [see Wolpin (2003)].

Lets us momentarily ignore investment and look solely at the implications of search. Consider a worker who receives a rental rate R_t for his human capital from his current employer in period t, so that $Y_t = KR_t$. Now imagine that during period t, the worker is matched with a new employer offering another rental rate, R. Because the worker can follow the same search strategy wherever he is employed, it is clear that the offer will be accepted if $R > R_t$ and rejected if $R < R_t$. If the worker rejects the offer and stays with the current employer, his earning capacity remains the same and $Y_{t+1} = Y_t$. If the

[5] The crucial feature here is that investment depends only the time left to the end of the horizon and is independent of the level of the stock of human capital. An additional simplification is that there is no depreciation, so that earning growth depends only on investment and is thus also independent of history. The results stated in the text can be easily shown using the continuous time version of the problem described in the text and applying phase diagram techniques [see Weiss (1986)].

[6] Weinberg (2003) shows that computer adoption is also related to experience. In industries with the greatest increase in computer use, the returns to experience have increased among high school graduates, but declined among college graduates.

worker accepts the outside offer and moves to the new employer, his new wage, $Y_{t+1} = RK$, must exceed Y_t. The probability that the worker will switch jobs is $\lambda(1 - F(R_t))$ and is *decreasing* in R_t.

The observable implications of this model are:

- A job has an *option value* to the worker. In particular, he can maintain his current wage and move away when he gets a better offer. Consequently, earnings rise whenever the worker switches jobs and remain constant otherwise.
- The higher the worker's current wage, the more valuable is the current job; hence, the offers that the workers accepts must exceed a higher reservation value. Therefore, the quit rate and the expected wage growth decline as the worker accumulates work experience and climbs up the occupational ladder.
- A straight-forward extension is to add involuntary separations. Such separations are usually associated with wage reduction and are more likely to occur at the end of the worker's career, which may explain the reduction in average wages towards the end of the life cycle.

This model can be generalized by allowing the worker to control the arrival of new job offers by spending time on the job in active search [see Mortensen (1986)]. Search effort declines as the worker obtains better jobs, so that the arrival rate of job offers and wage growth decline, too. Towards the end of the career, a worker may reduce his search effort to a level that generates no job offers. Consequently, voluntary quits and wage growth cease.

The same search model can be motivated slightly differently by assuming that workers and firms are heterogeneous. Let workers be ranked by their skill, K. Let firms be ranked by their *minimal* skill requirement R [see Weiss, Sauer and Gotlibovski (2003)]. Assume that worker K employed by firm R produces R if $K \geqslant R$ and 0 otherwise. Because workers with $K \geqslant R$ on job R produce the same amount, irrespective of their K, we can set their wages to R (assuming zero profits). A worker K who is now employed at firm R_t and meets (with probability λ) a random draw from the population of employers, R, is willing to switch if and only if $R > R_t$. However the employer is willing to accept him only if $K \geqslant R$. Transition into a better job thus occurs with probability $\lambda(F(K) - F(R_t))$.

3.4. *Comparison of investment and search*

The investment and search models have similar empirical implications for *average* growth in earnings, i.e., positive and declining wage growth. In the investment model, the reason for wage growth is that the worker chooses to spend some of his time learning. However, investment declines as a result of the shortened remaining work period, which causes wage growth to taper off. In the search model, wage growth is an outcome of the option that workers have to accept or reject job offers. Acceptance depends on the level of earnings that the worker attained by time t, so that history matters. Two workers of the same age may behave differently because of their different success records in meeting employers. But the general trend is for wage growth to decline because workers who attained a higher wage have a lower incentive to search and are less likely to switch jobs.

Although investment and search have similar implications for wage growth, they can be distinguished by their different patterns in the *variance* of wages and the correlation between wages at different points of the life cycle. As shown by Mincer (1974), the variance in wages first declines and then rises, as we move across age groups in a cross section or follow a cohort. The reason is that a current low wage is compensated for by a future high wage, so that workers who invest more intensely will *overtake* those with a lower investment rate. The minimal variance occurs in the middle range of experience, where individual earning profiles cross. Under search, the cause for variability is not differential investment but different success record in locating suitable job matches and the variability in accepted wage offers. Homogeneous workers become increasingly heterogeneous due to their longer exposure to random job offers. However, selection modifies the impact of such shocks on wages, because wages do not go down when the worker keeps the job and those who have high wages are less likely to get a better offer. Thus, the variance first increases and then declines as workers are gradually climbing up the income distribution. If workers are initially heterogeneous, the variance may also first increase and then decline as workers are gradually sorted into their "right" place. The investment model suggests a negative correlation between wage level and wage growth at the beginning of the worker's career and a positive correlation between wage growth and wage level late in the worker's career, whereas the search model implies a negative correlation between current wage and wage growth at *any* point of the life cycle.

Search and investment also have similar implications for quits, especially if investment has a firm-specific component. To the extent that specific investment can be described by a stochastic learning process on the job, as in Jovanovic (1984) and Mortensen (1988), then both wage growth and mobility can be outcomes of either internal shocks in the form of changes in the quality of a match, or external shocks in the form of outside offers. The average patterns of wage growth and separations will be the same under specific investment or search. However, higher moments, such as the wage variances among stayers and movers, can indicate the importance of specific capital and search, respectively.

3.5. Putting the two together

We now consider the possible interaction between search and investment behavior. To simplify, we continue to assume that workers can reject or accept offers as they arrive at an exogenous rate λ, but cannot initiate offers by investing in search. However, the option of passive search changes the incentives to invest in human capital.

The Bellman equation becomes

$$\begin{aligned} V_t(R_t, K_t) = \operatorname*{Max}_{l_t}\Big\{ & R_t K_t(1-l_t) + \beta\big[\lambda E\big\{\max\big[V_{t+1}(R_t, K_{t+1}), V_{t+1}(R, K_{t+1})\big]\big\} \\ & + (1-\lambda)V_{t+1}(R_t, K_{t+1})\big]\Big\}. \end{aligned} \tag{12}$$

Because a worker with a given K can follow the same search and investment strategy on any job, it is clear that he will switch jobs if $R > R_t$. Given this reservation value strategy, we can write

$$\begin{aligned} &E\big\{\max\big[V_{t+1}(R_t, K_{t+1}), V_{t+1}\big(\widetilde{R}_{t+1}, K_{t+1}\big)\big]\big\} \\ &\quad = F(R_t)V_{t+1}(R_t, K_{t+1}) + \int_{R_t}^{\infty} V_{t+1}(R, K_{t+1})f(R)\,\mathrm{d}R, \qquad (13) \end{aligned}$$

where $f(R)$ is the density of wage offers. The first-order condition for l_t is now

$$\begin{aligned} \frac{R_t}{g'(l_t K_t)} = \beta V_{k,t+1}(R_t, K_{t+1}) + \lambda\beta \int_{R_t}^{\infty} \big(&V_{k,t+1}(R, K_{t+1}) \\ &- V_{k,t+1}(R_t, K_{t+1})\big) f(R)\,\mathrm{d}R, \qquad (14) \end{aligned}$$

where $V_{k,t}$ denotes the partial derivative of $V_t(\cdot,\cdot)$ with respect to K_t. The interaction between investment and search decisions is captured by the second term in Equation (14) which shows that the incentives to invest now include the *capital gains* that the worker obtains if he changes employers. The higher K_t, the more one gains from a favorable draw of R; therefore, the incentive to accumulate human capital is stronger.[7]

This extended model has the following features:

- As long as the worker stays with the same firm, investment in human capital declines because of the shortened work period.
- On any such interval, the worker invests more than he would without search and a fixed R. This result reflects the upward drift in the R which is inherent in the search model and qualitatively similar to the result in the regular investment model when R rises exogenously.
- Investment drops when the worker switches to a new job with a higher R, because the option of switching to a new job becomes less valuable.

3.6. Human capital and skills

Human capital K is an aggregate that summarizes individual skills in terms of production capacity. Different skills are rewarded differentially in different occupations. We

[7] We can simplify these expressions by showing that the value function is *linear* in K_t and can be written in the form

$$V_t(R_t, K_t) = K_t A_t(R_t) + B_t.$$

Hence, investment in period t is determined by

$$\frac{R_t}{g'(l_t K_t)} = \beta A_{t+1}(R_t) + \lambda\beta \int_{R_t}^{\infty} (A_{t+1}(R) - A_{t+1}(R_t)) f(R)\,\mathrm{d}R,$$

where $A_t(x)$ is a sequence of functions that are increasing in x and decreasing in t, with $A_T(x) = 1$ for all x.

assume that this aggregate may be represented as

$$\ln K_j = \sum_s \theta_{sj} S_s, \tag{15}$$

where S_s is the quantity of skill s possessed by the individual and θ_{sj} is a non-negative parameter that represent the contribution of skill s to occupation j. Firms reward individual skills indirectly by renting human capital at the market-determined rental rate, R. Thus, the parameter θ_{sj} is the proportional increase in earning capacity associated with a unit increase in skill x_s if the individual works in occupation j. Having assumed that θ_{sj} is independent of the quantity of skill s possessed by the individual, these coefficients may be viewed as the implicit "prices" (or "rates of return") of skill s in occupation j.[8]

Because we are interested here in the timing of occupational changes, it will be convenient to set the problem in continuous time. We denote by T the duration of the worker's lifetime and by t a point in time in the interval $[0, T]$. We define $h_j(t)$ as the portion of available time spent working in occupation j at time t, so that $0 \leqslant h_j(t) \leqslant 1$ and $\sum_j h_j(t) = 1$. The worker will typically work at one particular occupation in each point in time but is free to switch occupations at any time. The worker's earning capacity is

$$Y(t) = R \sum_j h_j(t) K_j(t). \tag{16}$$

Skills are initially endowed and can then be augmented by acquiring experience. We consider here a "learning by doing" technology whereby work at a rate $h_j(t)$ in a particular occupation j augments skill s by $\gamma_{sj} h_j(t)$. Thus, the change in skill s at time t is

$$\dot{S}_s = \sum_j \gamma_{sj} h_j(t). \tag{17}$$

Note the joint production feature of this technology. Working in any one occupation j can influence many skills that are useful in other occupations. Yet, such experience may be more relevant to some particular skills. In this way, we obtain that work experience is transferable but not necessarily general.

In the static version of this model (the Roy model), individual skills are constant ($\gamma_{sj} = 0$ for all s and j) and the main issue is the mapping between skills and earnings that results from the different occupational choices of workers with different skills. The basic principle that applies there is that each individual will spend *all* his work time in the occupation in which his bundle of skills commands the highest reward [see Willis (1986) and Heckman and Honore (1990)]. Unexpected changes in the prices of skills, θ_{sj}, can cause the worker to switch occupations; however, under static conditions

[8] Each worker can be viewed as a bundle of skills. Because these skills cannot be unbundled from the worker, the law of one price does not apply and different skills have different implicit prices in different uses, depending upon the technology of production. It is only human capital that can be moved freely across uses and, therefore, commands a single price.

there is no occupational mobility. In the dynamic set up that we outline here, skills vary with time, and this variation is influenced by the worker's career choices. In such a context, *planned* occupational switches can arise, even in the absence of shocks, if experience is sufficiently transferable across occupations.

To simplify the exposition, we consider the case of two occupations and two skills[9] and examine the conditions for a single switch. Given our simplifying assumptions, the earnings capacity of a worker in different occupations, K_j grows at constant rates that depend on the occupation in which the worker specializes. Suppose that the worker switches from occupation 1 to occupation 2 at time x and then stays there for the rest of his life. Then, in the early phase, prior to time x, $h_1(t) = 1$ and

$$\begin{aligned}\frac{\dot{K}_1}{K_1} &= \theta_{11}\gamma_{11} + \theta_{21}\gamma_{21} \equiv g_{1,1},\\ \frac{\dot{K}_2}{K_2} &= \theta_{12}\gamma_{11} + \theta_{22}\gamma_{21} \equiv g_{2,1}.\end{aligned} \tag{18}$$

In the later phase, after x, $h_2(t) = 1$ and

$$\begin{aligned}\frac{\dot{K}_1}{K_1} &= \theta_{11}\gamma_{12} + \theta_{21}\gamma_{22} \equiv g_{1,2},\\ \frac{\dot{K}_2}{K_2} &= \theta_{12}\gamma_{12} + \theta_{22}\gamma_{22} \equiv g_{2,2}.\end{aligned} \tag{19}$$

The expected lifetime earnings of the worker is

$$V(x) = R\left\{K_1(0)\int_0^x \mathrm{e}^{-rt+g_{1,1}t}\,\mathrm{d}t + K_2(0)\int_x^T \mathrm{e}^{-rt+g_{2,1}x+g_{2,2}(t-x)}\,\mathrm{d}t\right\}. \tag{20}$$

For a switch at time x to be optimal, it is necessary that $V'(x) = 0$ and for $V''(x) < 0$. It can be shown that if work experience in each occupation raises the worker's earnings in that same occupation by more than in the alternative occupation (that is, $g_{1,1} > g_{2,1}$ and $g_{2,2} > g_{2,1}$) then $V'(x) = 0$ implies that $V''(x) > 0$, so that the worker will *never* switch occupations.[10] Instead, the worker will specialize in one occupation throughout his working life and concentrate all his investments in that occupation [see Weiss

[9] More generally, we are interested in the case where there are more occupations than skills. Otherwise some skills will be redundant [see Welch (1969)].

[10] The first derivative can be written in the form

$$\begin{aligned}V'(x) &= R\Big\{K_1(0)\,\mathrm{e}^{-rx+g_{1,1}x} - K_2(0)\,\mathrm{e}^{-rx+g_{2,1}x}\\ &\qquad + (g_{2,1}-g_{2,2})K_2(0)\int_x^T \mathrm{e}^{-rt+g_{2,1}x+g_{2,2}(t-x)}\,\mathrm{d}t\Big\}\\ &= R\,\mathrm{e}^{-rx+g_{2,1}x}K_2(0)\Big\{\mathrm{e}^{D+(g_{1,1}-g_{2,1})x} - 1 + (g_{2,1}-g_{2,2})\int_0^{T-x}\mathrm{e}^{-r\tau+g_{2,2}\tau}\,\mathrm{d}\tau\Big\}\end{aligned}$$

(1971)]. However, some occupations require a preparation period in other occupations, that serve as stepping stones [see Jovanovic and Nyarko (1997)]. For instance, it is not uncommon that successful managers start as engineers or physicians rather than junior managers.

Specifically, suppose that

$$\gamma_{11} > \gamma_{12}, \quad \gamma_{21} > \gamma_{22}, \quad \theta_{11} < \theta_{12}, \quad \theta_{21} < \theta_{22}. \tag{21}$$

Then it is easy to verify that, depending on initial conditions, the worker may start in occupation 1 and then switch to occupation 2 because skill 1 is more important in occupation 2, i.e., $\theta_{12} > \theta_{11}$, but occupation 1 is the better place to acquire skill 1, i.e., $\gamma_{11} > \gamma_{12}$. It does not pay to specialize in occupation 1 because the worker will not exploit his acquired skills that are more useful in occupation 2. Nor is it usually optimal to specialize in occupation 2, because then the worker will not acquire sufficient skills. However, a worker with a large endowment of skill 1 or skill 2 may specialize in occupation 2 immediately.

This model illustrates quite clearly the main features of occupations that serve as stepping stones. Basically, these occupations enable the worker to acquire skills that can be used later in other occupations in a cheaper or more effective way.[11] Although these jobs pay less for *all* workers with *given* skills, some workers may still enter them as an investment in training.[12]

The pattern of earnings growth that is implied by this sequence of occupational choices is easy to summarize. At the point of switch, x, earnings *rise* instantaneously, where the proportional jump is $S_1(0)(\theta_{11}-\theta_{12})+S_2(0)(\theta_{21}-\theta_{22})+(g_{1,1}-g_{2,1})x$. The growth rate of earnings may either rise or decline following this change, because the restrictions in (21) are consistent with either $g_{1,1} > g_{2,2}$ or $g_{1,1} < g_{2,2}$. If we assume, however, that the differences between the two occupations in the learning coefficients (the γ's) are more pronounced than the differences in the prices of skills (the θ's) then $g_{1,1} > g_{2,2}$ and the growth rate in earnings will decline, which is the more realistic case.

using $\frac{K_1(0)}{K_2(0)} = \mathrm{e}^{S_1(0)(\theta_{11}-\theta_{12})+S_2(0)(\theta_{21}-\theta_{22})} \equiv \mathrm{e}^D$ and a change of variable, $t - x = \tau$. The second derivative evaluated at this point is given by

$$V''(x) = R\,\mathrm{e}^{-rx+g_{2,1}x} K_2(0)\{(g_{1,1}-g_{2,1})\,\mathrm{e}^{D+(g_{1,1}-g_{2,1})x} - (g_{2,1}-g_{2,2})\,\mathrm{e}^{(-r+g_{2,2})(T-x)}\}.$$

[11] Because this model assumes learning by doing, the opportunity costs of investment are the forgone earnings that one could receive by switching earlier to the higher-paying occupation. This stands in some contrast to the cases discussed above, where the costs were the loss of effective work time, in the occupation that one has.

[12] Booth, Francesconi and Frank (2002), show that fixed term temporary jobs serve as stepping stones to permanent jobs. Female workers who held 3 consecutive one year fixed-term contracts are initially paid lower wages than comparable workers on permanent jobs but appear to overtake them after about 10 years. Among men, wages are higher for workers who follow the same pattern but overtaking is not observed.

3.7. Wages, productivity and contracts

The presumption, so far, was that a worker's wage is closely tied to his productivity. However, the relation between these two variables may be quite complex, especially when workers and firms develop durable relationships. In such a case, wages and productivity are still tied in terms of long-term averages but, in the short run, systematic differences between wages and productivity may appear that represent credit and risk sharing arrangements, or incentives to exert effort. We shall not attempt to describe the complex issues associated with incentives for effort, about which several excellent recent surveys exist. However, the issues associated with credit and risk sharing are easy to explain.

Trade between workers and employers that extends over time is motivated by some basic asymmetry between the parties. Specifically, firms may have better access to the capital market and may be able to pool some risks. If a worker's output varies over time, and if he has no access to the capital market, the firm may smooth his consumption by offering a flat wage profile which effectively means that the worker borrows from the firm. Similarly, if a worker's output is subject to shocks, the firm may accept these risks and provide the worker with insurance that stabilizes his income. As we shall now show, the ability of firms to provide such credit or insurance arrangements is limited by the commitments that workers (and firms) can make.

Consider a worker with a fixed bundle of skills and suppose that because of random variations in the prices of skills, his/her human capital is subject to capital gains or losses. Specifically,

$$K_{t+1} = \begin{cases} K_t(1+g) & \text{with probability } p, \\ K_t(1-\delta) & \text{with probability } 1-p, \end{cases} \tag{22}$$

where g and δ are fixed parameters that govern the size of capital gains and losses, respectively. We denote by $Q_t(K_{t-1})$ the expected present value of the worker's output over the remainder of his work life, $T-t$. Let h_t be a sequence of zeros and ones, where 1 for the τ element, $\tau = 1, 2, \dots, t$, indicates the occurrence of a positive shock and a 0 indicates the occurrence of a negative shock in period τ. We refer to such a sequence as the history or sample path. Let $y_t(h_{t-1})$ be the wage that a firm promises to pay a worker with history h_{t-1} in period t and let $Y_t(h_{t-1})$ be the present value of the expected payments over the remainder of the working life, from t to T.[13] We can think of $Y_t(h_{t-1})$ as the worker's *contractual* assets.

[13] For simplicity, assume that the interest rate is zero. Then, $Q_t(K_{t-1})$ satisfies the difference equation

$$Q_t(K_{t-1}) = RK_{t-1} + pQ_{t+1}(K_{t-1}(1+g)) + (1-p)Q_{t+1}(K_{t-1}(1-\delta)),$$

and can be solved recursively, using the end condition that $Q_{T+1}(K_T) = 0$ for all K_T. Similarly,

$$Y_t(h_{t-1}) = y_t(h_{t-1}) + pY_{t+1}(h_{t-1}, 1) + (1-p)Y_{t+1}(h_{t-1}, 0)),$$

can be solved recursively, using the end condition that $Y_{T+1}(h_T) = 0$ for all h_T.

A risk-neutral firm is indifferent between all contingent contracts that yield the same expected value. However, a risk-averse worker with no access to the capital or insurance markets would prefer that the payment stream will be as stable as possible. If the worker can commit to stay with the firm, the competition among firms will force them to offer wage contracts that smooth the wage payments over time and across states of nature. In practice, workers cannot legally bind themselves to a firm; their option to leave the firm limits the insurance and consumption smoothing that firms can provide [see Harris and Holmstrom (1982), Weiss (1984)].

A competitive payment scheme must maximize the expected utility of the worker given the firm's expected profits and the worker's outside options. Therefore, the contract that survives must solve the following program

$$V_t(K_{t-1}, Y_{t-1}) = \operatorname*{Max}_{y,x_1,x_0} \Big\{ (u(y) + pV_{t+1}\big(K_{t-1}(1+g), Y_{t-1}+x_1\big) \\ + (1-p)V_{t+1}\big(K_{t-1}(1-\delta), Y_{t-1}+x_0\big)\Big\}, \tag{23}$$

subject to

$$y + px_1 + (1-p)x_0 = 0, \tag{24a}$$

$$Y_{t-1} + x_1 \geqslant Q_{t-1}\big(K_{t-1}(1+g)\big) - a, \tag{24b}$$

$$Y_{t-1} + x_0 \geqslant Q_{t-1}\big(K_{t-1}(1-\delta)\big) - a, \tag{24c}$$

where a is a parameter that represents the costs of mobility across firms, such as loss of firm-specific capital.[14] The state variables at period t are the worker's human capital and the expected payments from the firm under the existing contract (including current obligations $y_t(h_{t-1})$). The control variables, y, x_1, x_0 represent possible revisions of that contract that can make the worker better off, keeping the firm's expected profits constant and keeping the worker with the firm.[15] Constraint (24a) requires that the revisions maintain the cost of the contract to the firm (because Q_{t-1} is fully determined by K_{t-1}, this implies that expected profits are unchanged). Constraints (24b) and (24c) imply that other firms cannot bid workers away. If the firm changes the contract in such a

[14] For simplicity, we treat the mobility cost as a fixed cost. In general, these costs depend on the time spent in a job, firm or industry and on the worker's particular skills.

[15] By definition,

$$\begin{aligned} Y_t(h_{t-1}) &= y_t(h_{t-1}) + pY_{t+1}(h_{t-1}, 1) + (1-p)Y_{t+1}(h_{t-1}, 0) \\ &= y + p(Y_t(h_{t-1}) + x_1) + (1-p)(Y_t(h_{t-1}) + x_0), \end{aligned}$$

implying

$$y + px_1 + (1-p)x_0 = 0.$$

manner that its obligation falls short of the worker's expected output, it cannot retain the worker because another firm can offer a superior contract and still make non-negative profits.

The first-order conditions are

$$u'(y) - \lambda = 0, \tag{25a}$$

$$\frac{\partial V_{t+1}(K_{t-1}(1+g), Y_{t-1} + x_1)}{\partial Y_t} - \lambda + \frac{\mu_1}{p} = 0, \tag{25b}$$

$$\frac{\partial V_{t+1}(K_{t-1}(1-\delta), Y_{t-1} + x_0)}{\partial Y_t} - \lambda + \frac{\mu_2}{1-p} = 0, \tag{25c}$$

where λ, μ_1, μ_2 are the time-variable non-negative Lagrange multipliers that are associated with the constraints (24a), (24b) and (24c), respectively.

Differentiating (23) with respect to Y_{t-1} and using conditions (25a)–(25c), we have

$$\frac{\partial V_t(K_{t-1}, Y_{t-1})}{\partial Y_{t-1}} = \lambda, \tag{26}$$

which implies that in each period and at any possible state, the marginal utility of consumption, $u'(y)$, is equated to the marginal value of the worker's contractual assets, $\frac{\partial V_t(K_{t-1}, Y_{t-1})}{\partial Y_{t-1}}$. Because the Lagrange multipliers μ_1 and μ_2 are non-negative, it follows from conditions (25b) and (25c) that the payment stream is arranged in such a way that the marginal value of contractual assets never rises. This also means that wage payments never decline as successive realizations of human capital unfold.

These results have a simple economic interpretation. Workers who may suffer either capital gains or capital losses, when skill prices change, would like the firm to transfer wages from "good" states when income is high and marginal utility of income is low to "bad" states when income is low and marginal utility of income is high. The firm is willing to do so only if the expected present value of wage payments does not rise in consequence. Thus, paying a higher current wage in a bad state implies a wage reduction in some future good state. However, the firm can commit to such a transfer policy only if it is able to retain the worker and collect the payment for the insurance that it provides the worker now.

If the cost of mobility across firms, a, is sufficiently high to prevent mobility, then constraints (24b)–(24c) are not binding and $\mu_1 = \mu_2 = 0$. Then, the optimal contract implies that y is a constant, which means that the firm provides *perfect* insurance and consumption smoothing. However, if the cost of mobility across firms, a, is sufficiently low, the constraint (24c) which corresponds to a positive shock is binding, because such a shock makes the worker more attractive to other firms. The wage profile that emerges in this case is one in which the wage rises when workers receive a positive shock but remains unchanged when they receive a negative shock. In this way, the workers receive *partial* insurance from the firm. When a positive shock occurs, wages are raised to the minimal level required to retain the worker. When a negative shock occurs, wages are set above the worker's productivity. This policy requires that workers pay for the insurance by accepting initial wages that fall short of their productivity upon joining the firm.

If the costs of mobility across firms are low, and workers must be induced to stay with the firm, then their *average* wages rise *faster* than their average productivity. This result is reversed if there are substantial costs of mobility across firms and the workers are locked to the firm, a condition that allows the firm to provide perfect insurance. In this case, of course, average wages rise at a *lower* rate than does productivity.

In equilibrium, there is no mobility across firms. However the workers' *option* to leave the firm affects wage growth. Paradoxically, workers are better off when the costs of mobility are high. This holds for two related reasons. First, with high mobility costs, workers are effectively locked in with the firm so that the firm can provide perfect rather than partial insurance. Second, because information is public and workers are equally productive in all firms, mobility serves no productive role. Thus the most efficient arrangement is for workers to stay with their employers. A more complex situation arises if workers can influence skill acquisition and use via occupational switches. Then, workers will receive less insurance from the firm but obtain higher wage growth resulting from investment in skills acquisition. In addition, workers may try to create a more balanced portfolio of skills, a factor supporting mobility and, possibly, multiple job holding.

An important feature of the optimal wage contract is that wages in period t generally depend on the entire history of shocks and not simply on the accumulated human capital at time t. Specifically, $y_t(h_{t-2}, 1, 0)$ may exceed $y_t(h_{t-2}, 0, 1)$. While workers have the same productive capacity in period t in both cases, there are wage gains from having *early* success. This is because early success provides opportunities for sharing risk with potentially more productive realizations in the future, an option not available to workers who experienced early failure. More generally, conditions at the time at which the commitments are taken, e.g., when workers entered the firm, can cause wage differences between workers who are equally productive.[16]

3.8. *Unobserved productivity and learning*

A particular worker's productivity may be unknown to the worker and potential employers. Over time, the worker's performance is observed; one may use this information to make inferences about the worker's "true" skills. This learning process can create negative and positive shocks to the worker's perceived productivity, similar to those discussed above. However, the learning model has further implications concerning mobility. That is, workers can experiment in an occupation where learning about ability is

[16] Two basic features of this model have been demonstrated empirically. First, nominal wages are indeed rigid downward [see Baker, Gibbs and Holmstrom (1994a, 1994b), McLaughlin (1994)]. However, the prevalence of real wage reduction is problematic for the contracting model. Second, history-dependence is in fact present [see Baker, Gibbs and Holmstrom (1994a), Beaudry and Di Nardo (1991)]. There is also evidence that risk aversion reduces wage growth [Shaw (1996)].

possible and then, as their abilities are gradually revealed, sort themselves into different occupations, based on their realized performance.

Let there be two occupations, one low skill, one high skill, and let there be two types of workers, those of high ability and those of low ability. All workers perform equally well in the low-skill occupation and produce one unit of output per period, irrespective of ability. Workers differ in their ability to perform the required jobs in the high-skill occupation; we denote the expected output, per period of time, as q_l and q_h for the low and high ability workers, respectively. However, neither the workers nor their employers know whether a particular worker is of high or low ability. The common prior probability that a specific worker is of low ability is denoted by π_0. With time, as a worker's performance is observed by all agents (including the worker himself), all agents modify this common prior.[17] Although a worker's productivity remains constant over time, the new information can affect his wages and employment.

We may model the realized output as a simple Bernoulli trials so that q_i is the fixed probability that type i, $i = l, h$, will produce one unit of output in period t and $1 - q_i$ is the probability that type i will produce nothing in period t. Let $n(t)$ be the (random) number of successes that a worker has accumulated up to period t. Based on this information, one can update the probability that he is of the low ability type. Specifically, the posterior probability is

$$\begin{aligned}\pi(t, r) &\equiv \Pr\{q = q_l / n(t) = r\} \\ &= \frac{\pi_0 q_l^r (1 - q_l)^{t-r}}{\pi_0 q_l^r (1 - q_l)^{t-r} + (1 - \pi_0) q_h^r (1 - q_h)^{t-r}},\end{aligned} \tag{27}$$

and the updated expected output per period is

$$q(t, r) = q_l \pi(t, r) + q_h \bigl[1 - \pi(t, r)\bigr]. \tag{28}$$

From (27) it follows that $\pi(t, r)$ rises in t for a given r and declines with r for a given t. That is, if a worker did not perform well, a low $n(t)$ up to a given time t, the posterior probability that he is of low ability increases. In contrast, if the worker has a favorable record, the posterior probability that he is of high ability increases. The perceived (expected) output of the worker is correspondingly modified downwards or upwards. (In this respect, the model is similar to the one discussed in the previous section, except that the informational value of the shocks (success or failure) decays over time.) With sufficient time, the process reveals the true identity of the worker.[18]

[17] We examine here only learning that is general for all firms in a particular industry. As already noted, firm-specific learning, involves some complex issues about the nature of the competition among firms that we cannot cover here. See, however, Jovanovic (1979a, 1979b, 1984), Mortensen (1988), Felli and Harris (1996, 2003) and Munasinghe (2003).

[18] Rewrite

$$\pi(t, r) = \frac{1}{1 + \frac{1-\pi_0}{\pi_0} (\frac{q_h}{q_l})^r (\frac{1-q_h}{1-q_l})^{t-r}}.$$

Consider first the case in which workers are risk-neutral and assume that workers are paid their current perceived output at each point of time. Because all workers are ex ante identical, they will all start at the risky high skill occupation, while attempting to learn their true ability. As the public information about each worker accumulates, workers are separated in terms of wages and employment. Those with inferior performance will receive lower wages and some of them will choose to leave. Those with superior records will receive higher wages and will choose to stay. Because of the finite time horizon and costs of mobility, workers will not move at the end of their career even though their perceived output and wages continue to fluctuate. This mobility pattern continues to hold if workers are risk-averse and if firms provide partial insurance so that wages are rigid downwards. However, an important difference is that such insurance can induce the workers to stay in the skilled sector even if their output in that occupation is low. With efficient contracts, such workers must be forced out, i.e., denied tenure [see Harris and Weiss (1984)].

The "pure" learning model has some strong implications for wage growth that hold for any distribution of shocks provided that we continue to assume that the shocks are independent across time. Suppose that worker i's performance in period t is given by

$$y_{it} = \eta_i + \varepsilon_{it}, \tag{29}$$

where η_i is a fixed parameter that is unknown to the firm, and ε_{it} is a random i.i.d. shock with zero mean. Now if firms pay wages based on workers perceived output at time t, $w_{it} = E(y_{it}/I_t) = E(\eta_i/I_t)$, where I_t is any information available at t. Then, because expectations are linear operators, it follows that $E(\eta_i/I_t) = E(E(\eta_i/I_{t+1})/I_t)$ and

$$w_{it} = E(w_{i,t+1}/w_{it}). \tag{30}$$

This martingale property implies that innovations in the wage process $w_{i,t+1} - E(w_{i,t+1}/I_t) = w_{i,t+1} - w_{it}$ are serially uncorrelated. Intuitively, any particular piece of the agents' information that the researcher observes has already been used by the agents and cannot change the predicted outcome [see Farber and Gibbons (1996)]. However, if one adds contracting and downward rigidity due to risk aversion, then, conditioned on the current wage, history matters. In particular, early success implies higher wages throughout the worker's career. Nevertheless, if a person with an early success is compared to a person with a late success, but both receive the *same current wage* then the late beginner will have the higher future expected wage [see Chiappori, Salenei and Valentin (1999)]. That is, the fact that the early beginner has the same wage as a late beginner speaks against him. In this respect, "what have you done for us lately" matters more.

Farber and Gibbons (1996) and Altonji and Pierret (2001) discuss further empirical implications of such models of public learning. Importantly, they distinguish between

Then, holding r fixed, $\pi(t,r)$ approaches 1 and $q(t,r)$ approaches q_l as t rises. Similarly, holding $t-r$ constant, $\pi(t,r)$ approaches 0 and $q(t,r)$ approaches q_h as t and r rise together.

information available to an outside observer (econometrician) and the information available to the economic agents. If the econometrician can observe a variable that is correlated with ability, even if not observed by the agents, then this variable will have an affect on wages which rises with time, reflecting the accumulation of information by the agents. In contrast, the effects of outcomes that employers observe, other than the worker's output, and that are correlated with ability (such as schooling) will decline over time as their marginal informational content diminishes.

4. Basic findings and their interpretation

In this section we provide a second look at the data, while stressing findings that have some bearing on the alternative models of wage growth.

4.1. Mincer's earnings function

Jacob Mincer discovered an important empirical regularity in the wage (earnings) structure. Average earnings of workers (in a given schooling-experience group) are tied to schooling and work experience in a relatively precise manner as summarized by the now familiar Mincer equation

$$\ln Y_{it} = \alpha + \beta s_i + \gamma (t - s_i) - \delta (t - s_i)^2 + \cdots, \qquad (31)$$

where Y_{it} are annual earnings (or weekly or hourly wage) of person i in year t, s_i are the years of schooling completed by person i and $(t - s_i)$ are the accumulated years of (potential) work experience of person i by year t.

In his 1974 book, Mincer estimated this specification for a sample of about 30.000 employed males taken from the US 1960 census; he reported a coefficient of .107 for schooling and .081 and $-.0012$ for the two experience coefficients. Including weeks worked as an explanatory variable, the effects of experience declined to .068 and $-.0009$, implying that wages grow less than earnings. The same equation has since been estimated in many countries for different periods and sectors, with similar results.[19]

Mincer's important insight was that this stability is no accident but rather a reflection of powerful and persisting economic forces. In an early (1958, pp. 284–5) paper, he wrote that:

> The starting point of an economic analysis of personal income distribution must be an exploration of the implications of the theory of rational choice. An implication of rational choice is the formation of income differences that are required

[19] Mincer has estimated several variants of this equation. Apart from alternative time shapes for the experience profiles, he was also concerned about whether schooling has a diminishing impact, the interaction between schooling and experience and the role of labor supply. These are empirically important issues yet the version in the text has become most popular in subsequent applications.

> to compensate for various advantages and disadvantages attached to the receipts of incomes.... This principle, so eloquently stated by Adam Smith has become a common place in economics. What follows is an attempt to cast one important aspect of this compensation principle into an operational model that provides insights into some features of the aggregative income distribution and into a number of decompositions of it which recent empirical research has made possible. The aspect chosen concerns differences in training among members of the labor force.

To apply the compensation principle to the data, Mincer considered long-lived individuals who operate in a stationary economy with access to a capital market and maximize the present value of their lifetime incomes. Suppose that the different occupations (jobs) pay wages that depend on the worker's schooling and experience and can be described by some earnings (wage) function of the form $Y_j(s, t-s)$. Given that workers can choose schooling and then occupations (jobs) that require different levels of training, what form should these functions have in equilibrium? One basic condition is that the present value of different lifetime earnings streams must be equal. Otherwise, all workers will be attracted to the highest paying j, s option, and no one will choose any other option. This condition alone puts strong restrictions on the equilibrium wage structure and, in particular, it implies that the marginal contribution of schooling is the *same* for all occupations, irrespective of the time shape of the experience profile, which is a form of separability. A simple functional form that satisfies these requirements for a large T is $Y_j(s, t-s) = \mathrm{e}^{rs} y_j(t-s)$, where $\int_0^\infty \mathrm{e}^{-r\tau} y_j(\tau)\,\mathrm{d}\tau$ is a constant that is independent of j.[20] Taking logs, one gets that

$$\log Y_j(s, t-s) = y_0 + rs + \log y^e(t-s) + \varepsilon_{tj}, \tag{32}$$

where $y^e(t-s_j)$ is the mean effect of experience and $\varepsilon_{tj} = \log y^e_j(t-s) - \log y^e(t-s)$ are deviations caused by differences in on the job training across occupations.

This simple model highlights several general points:

- The effect of schooling on the log of wages is determined by the prevailing interest rate, reflecting the delay in receiving income that is implied by investment in schooling. Under this interpretation, it is important that schooling be measured in

[20] Letting $T = \infty$ and writing

$$V_j(s) = \int_s^\infty \mathrm{e}^{-rt} Y_j(s, t-s)\,\mathrm{d}t = \mathrm{e}^{-rs} \int_0^\infty \mathrm{e}^{-r\tau} Y_j(s, \tau)\,\mathrm{d}\tau,$$

we see that

$$V_j'(s) = -rV_j(s) + \mathrm{e}^{-rs} \int_0^\infty \mathrm{e}^{-r\tau} \frac{\partial}{\partial s} Y_j(s, \tau)\,\mathrm{d}\tau.$$

Thus, the conditions that $V_j(s)$ is a constant for all s and j and that $V_j'(s) = 0$ together imply that $\mathrm{e}^{-rs} \int_0^\infty \mathrm{e}^{-r\tau} \frac{\partial}{\partial s} Y_j(s, \tau)\,\mathrm{d}\tau$ is a constant for all s and j. Proceeding in this fashion, we obtain similar conditions for all higher-order derivatives. The specification in the text satisfies all these requirements, provided that $\int_0^\infty \mathrm{e}^{-r\tau} y_j(\tau)\,\mathrm{d}t$ is independent of j.

years. Moreover, if workers care only about income and leisure has little value, earnings rather than hourly (or weekly) wages should be the dependent variable.

- The average log earnings profiles of workers with different schooling are parallel, reflecting the separability of investment decisions in school and on the job.
- Individual earnings profiles intersect because they must provide the same present value of lifetime earnings. To the extent that a common effect for experience is used to describe earnings, the errors must be correlated over the life cycle so that early negative residuals imply positive late residuals and the variance of these residuals must be a U-shaped as a function of experience.
- These features are independent of demand conditions and should hold as long as individuals are homogeneous and schooling and occupations can be freely chosen, without barriers to entry. Significantly, these features may hold in different countries or periods, with different technologies and different demands for educated workers. In this respect, the model is classical. Prices are determined by an infinitely elastic supply and demand determines only the number of workers of each type.

Mincer then used Becker's 1967 Woytinsky lecture [reprinted in Becker (1975)] and Ben-Porath's (1967) results on optimal investment in human capital to put restrictions on the average contribution of experience to earnings, $y^e(t-s)$. He notes that: "learning from experience is an investment in the same sense as the more obvious forms of on-the-job training, such as, say, apprenticeship programs. Put in simple terms, an individual takes a job with an initially lower pay than he could otherwise get because he knows that he will benefit from the experience gained in the job taken" (1993, vol. 1, p. 102). He then notes that: "Generally speaking, the fact that age-earnings profiles slope upward over part of the life cycle is a consequence of the tendency to invest in human capital at young ages. ... Investments are spread over time because the marginal costs of producing them is upward sloping in each period. They decline over time because marginal benefit decline and because the marginal cost curve shifts upward" (1993, vol. 1, p. 44). The decline in benefits reflects the fact that one can only exploit human capital by "renting" it out, but not by selling it. The increase in costs reflects the fact that investment in human capital requires the person's own time which is diverted from work.[21]

Let $k(t) = \frac{Y(t)}{K(t)}$ denote the portion of earning capacity that is utilized in the form of actual earnings; then, by definition, $Y(t) = K(t)(1 - l(t))$. Assume that $\frac{\dot{K}(t)}{K(t)} = rl(t)$ and that the investment ratio $l(t)$ equals 1 during schooling and then declines linearly with experience during the work period, i.e., $l(t) = a - b(t - s)$ for $t \geqslant s$. One thus obtains

$$\ln Y(s, t - s) \cong \ln K(0) + rs + r \int_0^{t-s} (a - bx)\,\mathrm{d}x - \big(a - b(t - s)\big), \tag{33}$$

which has the same functional form as the earnings function specified in (31).

[21] These two features are the main differences between the theories of investment in human and physical capital.

It his 1974 book, Mincer used these considerations to provide a direct economic interpretation for the coefficients of his estimated "human capital earnings function". The estimated coefficient on schooling in equation (30) reflects "the rate of return for schooling" and the coefficients on experience reflect the shape of the average person's investment profile. The reduction in investment is thereby tied to the observed slope and concavity of log earnings-experience profiles.[22]

As pointed out by Rosen (1977), under the model's strict assumptions, in particular the assumption that all earnings profiles yield the same present value, the life cycle pattern of earnings is undetermined. Thus, to use the human capital model, one must specify a particular trade-off between current and future earnings, usually called the "production function" of human capital. Thus, let $\dot{K} = g(I)$, where $I = lK$ and $g(I)$ is rising and concave. The assumptions that $g(I)$ rises and Y declines in I maintain the idea of compensation because one must sacrifice current earnings in order to increase earning capacity (and future earnings). The added assumption of concavity can be justified by the fact that a person must use his own resources to augment his earning capacity. But this would force identical individuals to choose the same investment path on the job. Differences in individual earnings profiles cannot, then, be simply attributed to differences in investments; individual attributes such as ability or access to the capital market, which affect individual "propensity to invest", must be introduced. In this case, it is no longer true that, in equilibrium, all income profiles are equivalent and that the observed wage ratios are independent of demand.

Mincer has often relied on Becker's (1975) analysis (first presented in his 1967 Woytinsky lectures) of the roles of ability and access to the capital market as factors affecting individual differences in investment. He is quite explicit in stating that: "Once ability and opportunity are introduced as determinants of investment, earning differentials can no longer be considered as wholly compensatory. Rents or "profits" from investment in human capital arise..." (1993, vol. 1, p. 59). These rents depend on the individual's attributes and on how much he chooses to invest. Mincer thus often refers to the estimated returns for schooling and experience as average returns.

Nevertheless, the role of individual heterogeneity initiated a major debate about the economic interpretation of the coefficients in the Mincer earnings functions. Given that these rates are based on comparison of different individuals who choose different levels of schooling, the casual effect of schooling is not identified, because it may simply reflect the impact of omitted (unobserved) ability and the positive correlation between ability and schooling [Griliches (1977)]. This debate was further stimulated by theoretical criticisms, based on asymmetric information and signaling, showing that schooling may have a positive effect even if it has no impact on a worker's output. More generally, to the extent that schooling is mainly a sorting device, social rates of return may be far lower than the private returns captured in the cross section.

[22] It is, of course, not necessary to assume investment in human capital to obtain such results. Rising and concave earning profiles can be also motivated by various forms of selection, such as the dismissal of unsatisfactory workers [see Flinn (1997)].

Huge research effort, based on twin data, natural experiments, and using variety of instrumental variables methods has tried to identify the causal effect of schooling. These studies generally follow Becker's scheme and assume that the individual level of schooling is determined by equating the marginal lifetime benefits of schooling with the marginal costs of financing it. The object of interest in these studies is the expected increase in average annual log earnings if a random sample (in a particular population) were to acquire an additional unit of education. The same interpretation of the rate of return holds in Mincer's compensating differences model, applied on the individual level. A person who is arbitrarily moved to a schooling program that requires one additional year of study will have proportionally higher future annual earnings (and output) given by the common interest rate, although there is no gain in lifetime earnings (or output). The crucial difference is that Mincer provides a market level analysis in which the contribution of schooling to earning is determined rather than taken as given. It is quite amazing that, after all this work, it was found that the impact of ability on the estimated rates of return is apparently not large and that Mincer's estimates of the average rates of return to schooling survived unscathed [see Card (1999, 2001)].

It must be recognized, however, that individual differences in ability can change the equilibrium structure in a fundamental way. The supply of workers of different skills is now positively sloped and the slope depends on the distribution of ability in the population. In this case, the rate of return to schooling depends on demand conditions. In addition, workers with different abilities invest differentially and have different lifetime earnings. Only "marginal" workers receive compensation for their investment, while other workers obtain ability rents. Further complications arise if ability is not unidimensional, and different workers fit different jobs [as in Willis and Rosen (1979)], or if ability is not observed by employers [as in Altonji and Pierret (2001)].

Similar problems arise with respect to the estimated impact of work experience on wage growth: Can we interpret the estimated coefficients of experience in Mincer's equation as the causal impact of investment on the job, or are they severely contaminated by differences in the attributes of the individuals choosing different levels of investment on the job? Moreover, how is trade off between current and future incomes determined in equilibrium? These issues are more difficult to resolve in the case of post schooling investments because the observed outcome is a whole wage profile rather than a single wage level and because, in contrast to schooling, investment on the job is not observed. Nevertheless, using panel data, one may examine properties of individual life cycle profiles to tease out some qualitative answers.

4.2. The variance covariance structure of earnings

One of Mincer's (1974) important findings is that the variance of the residuals from his estimated wage function forms a U-shaped function of potential work experience. This finding is quite surprising given that alternative models of life cycle earnings, such as learning or search, predict a monotonically increasing variance or a variance that is first increasing and then decreasing.[23] Mincer has interpreted this result as a consequence of compensating wage differences. That is, individual variation in the "propensity to invest" generates substantial differences at the early and the late stages of the life cycle, when workers who choose to invest first pay for their training and later receive the benefits. Mincer (1974) provides evidence supporting his U-shape prediction. Again, Mincer's early findings appear surprisingly robust. Heckman, Lochner and Todd (2001) confirmed Mincer's findings using later data and Polachek (2003) brought evidence for such patterns across countries. Figures 6a to 6e show the gap in log wages between the 90th and 10th percentiles within the education and experience categories, using the CPS repeated cross-sectional data for the periods 1964–1979 and 1980–2001. Like Mincer (1974) and Heckman, Lochner and Todd (2001), we find that the interpersonal wage dispersion exhibits a U-shape pattern, which is less pronounced at higher levels of schooling. As in Polachek (2003), we find that in recent years, the "break-even point" at which the variance is at its minimum (i.e., the experience level at which the earnings of investors and non-investors coincide) appears quite early in a career, approximately 3 to 5 years after entry into the labor market. The higher variability in the second period, 1980–2001, reflects the general increase in wage inequality due to changing skill prices. Nevertheless, the U-shape pattern persists in both periods.

The PSID and NLSY panels are too small to provide reliable estimates of the experience (time) patterns of the variance within education cells. We therefore follow Mincer (1974) and examine the variance of the residuals from a log wage regression equation that is linear in school years and quadratic in experience. For both panels, we obtain that the variance rises with labor market experience. It is only when we add individual fixed effects and consider the deviations for each person around the individual mean (over all the years that person was observed working) as well as the average wage profile of the sample that the U-shape pattern for the residual variance emerges (see Figures 7a and 7b).[24] Moreover, the minimum variance in both panels occurs at about ten years of experience, which is very close to Mincer's theoretical prediction. This suggests the presence of heterogeneity, meaning that individuals who invest more also have higher potential wages in the absence of investment. To address this possibility, one must go beyond the comparisons of different individuals, observed at different points of

[23] In learning models, the current wage is determined by a sum of independent shocks, which has an increasing variance, although the growth rate of this variance may decline as individuals are gradually sorted out. In the search model, job offers arrive randomly and independently over time, which causes an initial increase in the variance among identical workers with the same initial wage. However, selection modifies the impact of such shocks on wages, because wages cannot go down and those who have high wages are less likely to get a better offer.

[24] This result is quite different from that obtained by comparing the wage variance profiles of workers with the same life time earnings, where monotonicity of the wage profiles is sufficient to generate the U-shape pattern. By eliminating the average wage profile, the residuals reflect individual deviations from the general time pattern. By taking away the individual mean, we impose that all residual profiles sum to zero for each person and, therefore, must cross each other. However, we do not impose any specific time pattern. Importantly, the residual profiles need not be monotone, although the earning profiles generally are.

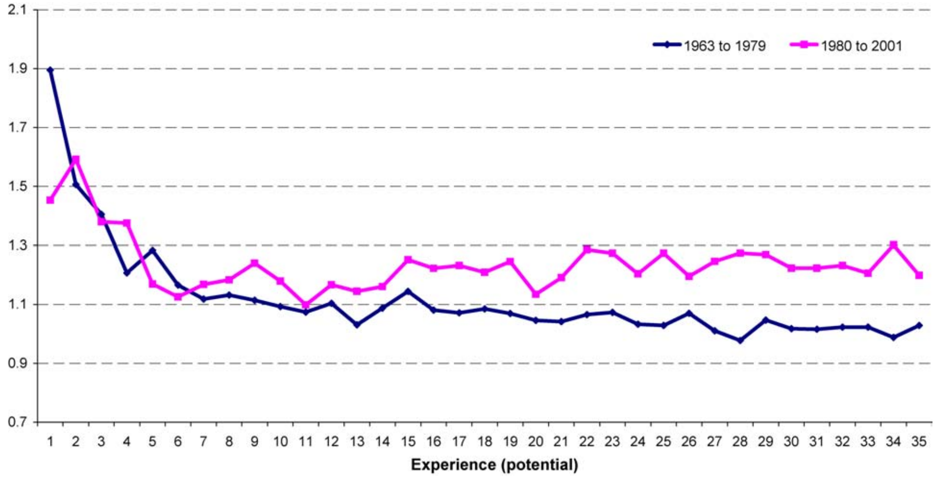

Figure 6a. The gap between workers belonging to the 90'th and 10'th percentiles of the residual log wage distribution for the periods 1963–1979 and 1980–2001, by education and experience, March CPS supplements, 1964 to 2002. High school dropouts.

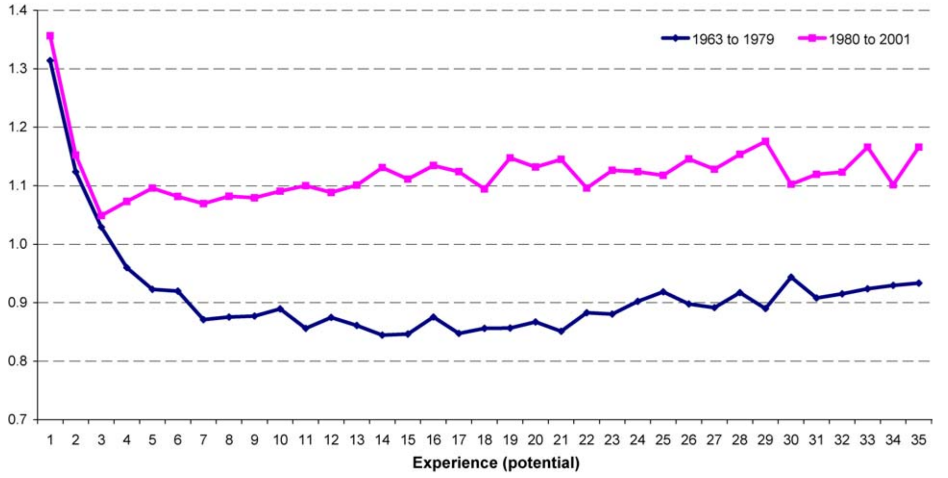

Figure 6b. The gap between workers belonging to the 90'th and 10'th percentiles of the residual log wage distribution for the periods 1963–1979 and 1980–2001, by education and experience, March CPS supplements, 1964 to 2002. High school graduates.

their career, and examine the properties of individual life cycle profiles by using panel data.

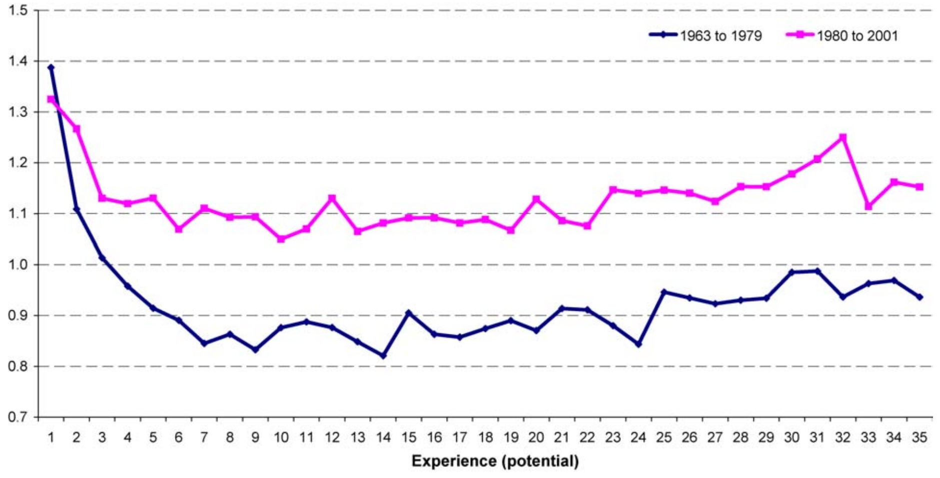

Figure 6c. The gap between workers belonging to the 90'th and 10'th percentiles of the residual log wage distribution for the periods 1963–1979 and 1980–2001, by education and experience, March CPS supplements, 1964 to 2002. Some college.

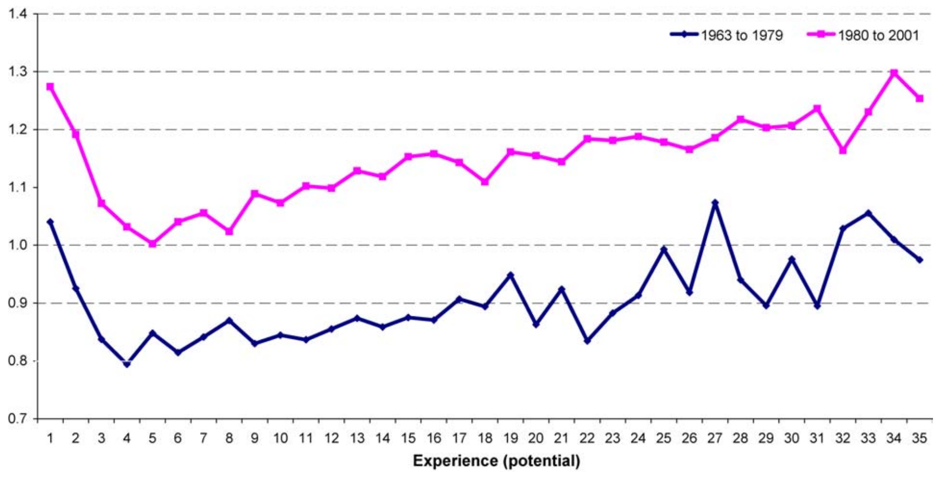

Figure 6d. The gap between workers belonging to the 90'th and 10'th percentiles of the residual log wage distribution for the periods 1963–1979 and 1980–2001, by education and experience, March CPS supplements, 1964 to 2002. College graduates.

In Figures 8a–8e, we take a first glance at the correlations between wage growth and wage level. The figures show the estimated coefficients and confidence intervals from a regression of wage growth on prior wage level by experience and education. To reduce the role of measurement errors, we look at three-year averages of these variables.

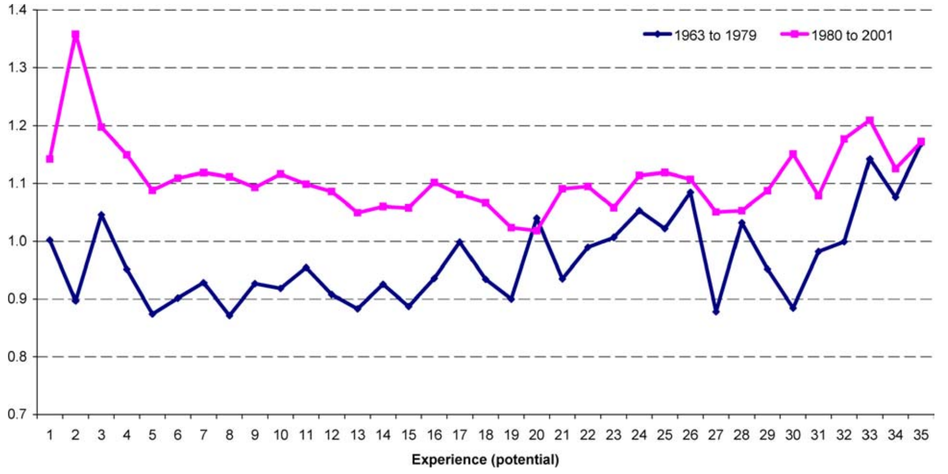

Figure 6e. The gap between workers belonging to the 90'th and 10'th percentiles of the residual log wage distribution for the periods 1963–1979 and 1980–2001, by education and experience, March CPS supplements, 1964 to 2002. Advanced degrees.

We see that within each experience group, there is a *negative* correlation between the current wage level and subsequent wage growth. This pattern is consistent with search behavior, because high-wage individuals are less likely to obtain superior offers. The investment model would suggest that the correlation is initially negative because low wages imply high investment, but later becomes positive as the high investment results in overtaking. In contrast, we observe negative correlations in all years. Yet, the fact that the correlations weaken as we move to higher experience groups suggests a presence of investment considerations.

To further examine the role of investment, we take a closer look at the covariance between earning levels at different points of time. The correlation matrices in Table 3 display the correlations between wages (and residuals obtained from the estimated Mincer wage equation, with and without individual fixed effects) at different stages of the life cycle. We use a balanced panel from the NLSY, where we again take three year averages. The correlation between income levels at different stages of the life cycle decays with the time distance, but is always positive. This result holds true also when we take residuals, eliminating the effects of schooling and experience. It is only when we eliminate the fixed effect of each person and consider the residual variation around the individual means (over all time periods) and the group average wage growth that we find negative correlations between early and late residuals. Moreover, these correlations become more negative as the time distance increases, providing clear evidence for compensation, whereby an early wage that is below the individual mean is associated with a late wage that is above the individual mean.

Thus, to identify compensation one must eliminate heterogeneity among individuals. Obviously, if individuals differ permanently in their earning capacity a positive corre-

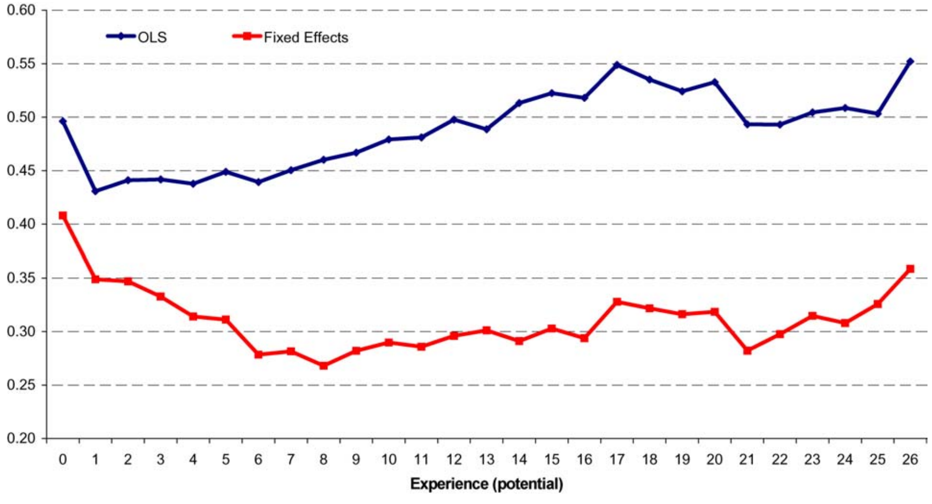

Figure 7a. Standard deviation of residuals in Mincer's wage equation by experience, with and without fixed effects, PSID, 1968–1997 and NLSY, 1979–2000. PSID, 1968–1997.

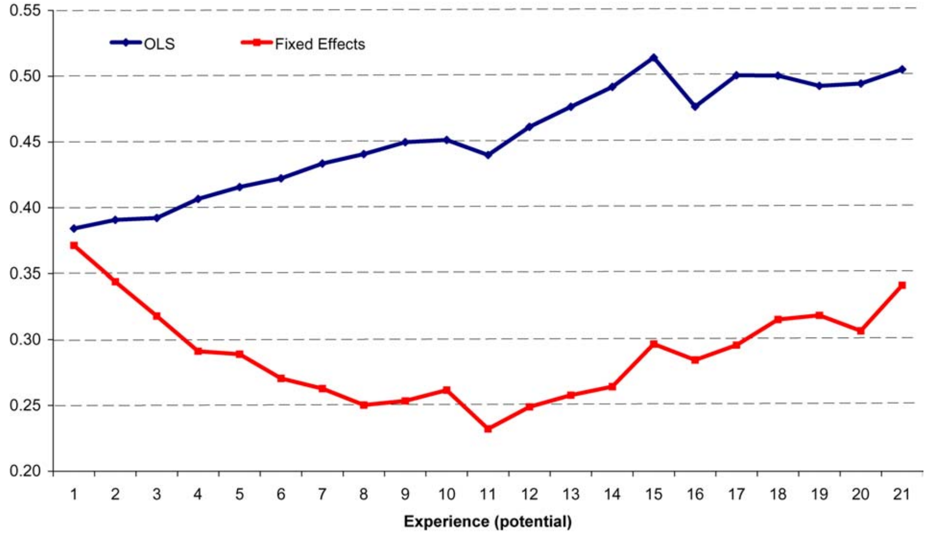

Figure 7b. Standard deviation of residuals in Mincer's wage equation by experience, with and without fixed effects, PSID, 1968–1997 and NLSY, 1979–2000. NLSY, 1979–2000.

lation will exist between early and late wages within each cohort because individuals who are above the mean are likely to remain above the mean, irrespective of investment. However, there may be more complex forms of heterogeneity that interact with experi-

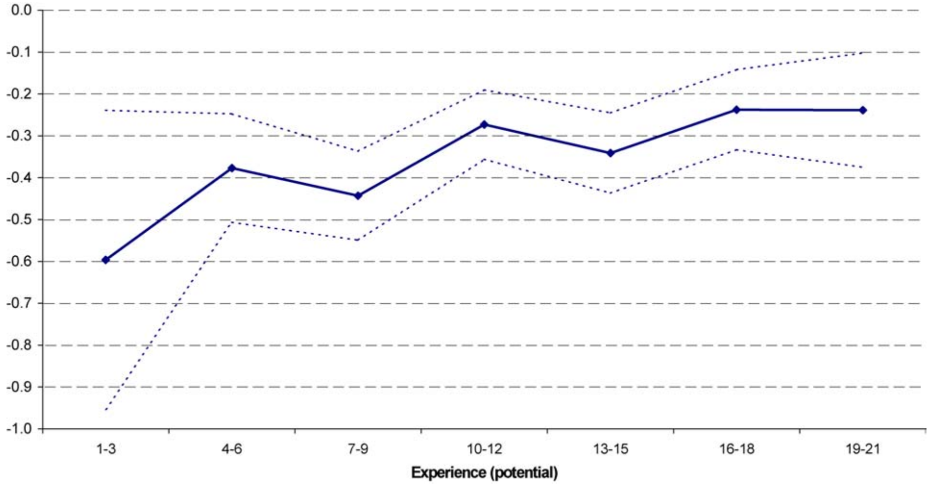

Figure 8a. Regression coefficients and confidence intervals of annual hourly wage growth rates on log hourly wage levels in prior period (3 year averages), by experience and schooling, NLSY, 1979–2000. High school dropouts.

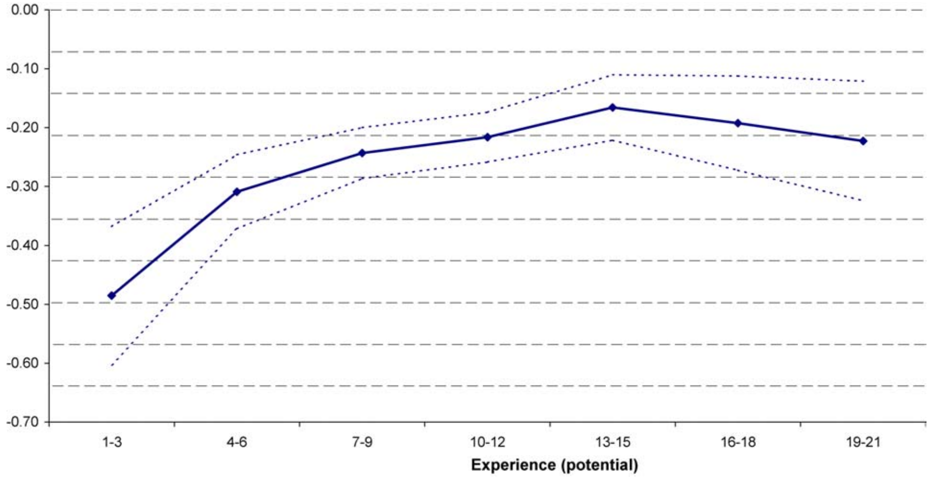

Figure 8b. Regression coefficients and confidence intervals of annual hourly wage growth rates on log hourly wage levels in prior period (3 year averages), by experience and schooling, NLSY, 1979–2000. High school graduates.

ence. In particular, there may be "systematic heterogeneity", whereby individuals with higher initial earning capacity also tend to invest more.[25] As explained in Mincer (1974,

[25] Note that initial earnings understate the individual's initial earning capacity and the bias depends on the propensity to invest. Mincer proposed to estimate initial earning capacity by the level of earnings at the "break even point", which he estimated to be about 10 years of work experience.

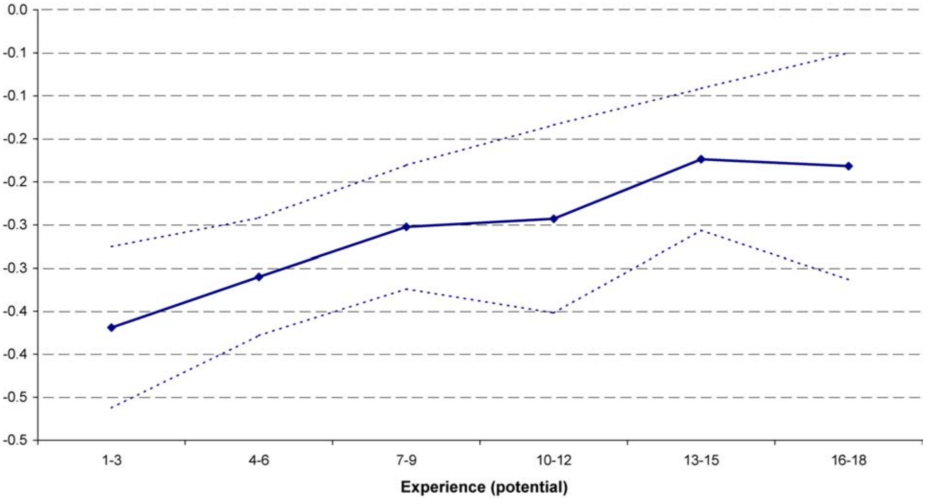

Figure 8c. Regression coefficients and confidence intervals of annual hourly wage growth rates on log hourly wage levels in prior period (3 year averages), by experience and schooling, NLSY, 1979–2000. Some college.

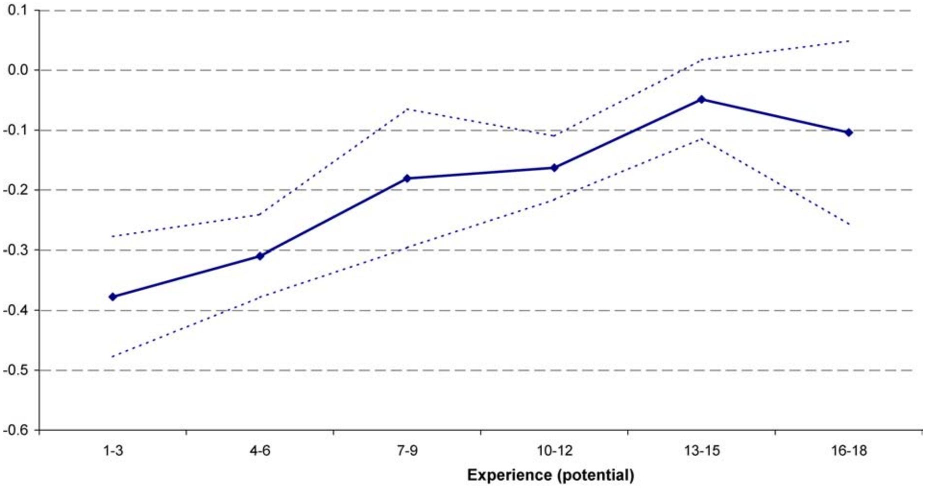

Figure 8d. Regression coefficients and confidence intervals of annual hourly wage growth rates on log hourly wage levels in prior period (3 year averages), by experience and schooling, NLSY, 1979–2000. College graduates.

Ch. 2) such heterogeneity tends to raise the within-cohort variance in earnings with the passage of time and may offset the effects of compensation.

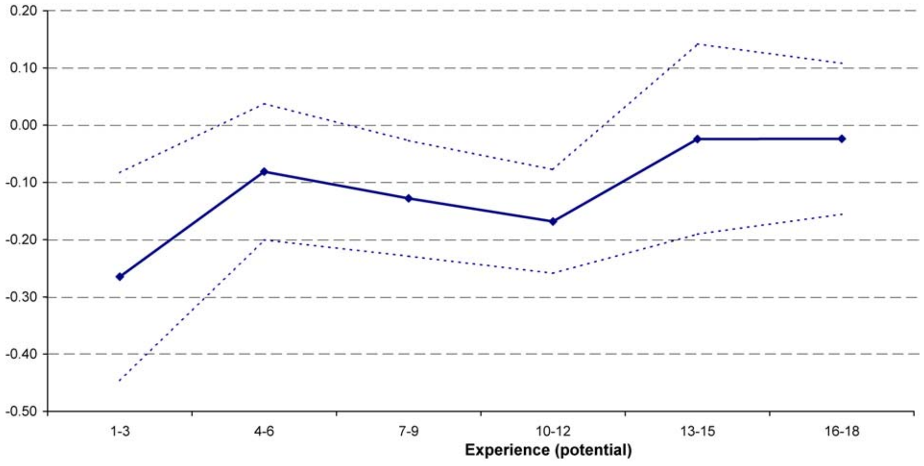

Figure 8e. Regression coefficients and confidence intervals of annual hourly wage growth rates on log hourly wage levels in prior period (3 year averages), by experience and schooling, NLSY, 1979–2000. Advanced degree.

Figure 9a displays estimated coefficients from regressions of individual fixed growth effects on individual fixed level effects, where the level effects are evaluated at two different points in the life cycle. When the level effect is the usual individual fixed effect, i.e., the mean wage residual during an individual career, the relationships between level and growth are significantly positive in all schooling groups but stronger among the highly educated. In such a case, we can interpret the level as a proxy for the individual's initial earning *capacity* and can conclude that individuals with higher "ability to learn" also have higher "ability to earn". However, if one evaluates the fixed effect as the intercept of the individual residual profile at the beginning of the worker's career, the relation becomes negative. In this case, the level effect also reflects investment, and the negative correlation reflects the fact that individuals with a higher propensity to invest forego a larger proportion of their initial earning capacity.[26]

In Figure 9b we present the regression coefficients of the individual slope and level (evaluated at the mean) on AFQT, which is an observable measure of individual ability. We see that *both* the level and growth effects are positively correlated with AFQT, which supports our interpretation of the previous results whereby individuals with higher "ability to learn" also have higher "ability to earn".[27] However, we do not find strong

[26] Baker (1997) and Haider (2001) report a *negative* correlation between the individual slopes and intercepts that evaluate an individual's deviation from the mean at *zero* experience. In contrast, Lillard and Weiss (1979) report a *positive* correlation between individual slopes and the mean residual (averaged over all experience levels). These findings are not inconsistent and indicate the presence of both heterogeneity and compensation.

[27] Although the impact of AFQT on the slope is significant only among the highly educated, it becomes significantly positive when we control for the initial level of the individual intercept. This suggests that the

Table 3
Correlations of log wages and residuals at different stages of the life cycle (three-year averages), full-time workers, NLSY, 1979–2000

(i): Log wage levels

Experience	1–3	4–6	7–9	10–12	13–15	16–18	19–21
1–3	0.195						
4–6	0.606	0.173					
	(0.000)						
7–9	0.476	0.738	0.193				
	(0.000)	(0.000)					
10–12	0.424	0.646	0.817	0.211			
	(0.000)	(0.000)	(0.000)				
13–15	0.374	0.588	0.701	0.789	0.238		
	(0.000)	(0.000)	(0.000)	(0.000)			
16–18	0.314	0.533	0.643	0.691	0.789	0.271	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
19–21	0.321	0.531	0.629	0.673	0.740	0.783	0.300
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

(ii): Residuals of Mincer's wage function

Experience	1–3	4–6	7–9	10–12	13–15	16–18	19–21
1–3	0.181						
4–6	0.563	0.151					
	(0.000)						
7–9	0.415	0.698	0.166				
	(0.000)	(0.000)					
10–12	0.358	0.592	0.788	0.183			
	(0.000)	(0.000)	(0.000)				
13–15	0.297	0.522	0.653	0.755	0.206		
	(0.000)	(0.000)	(0.000)	(0.000)			
16–18	0.230	0.459	0.586	0.644	0.757	0.236	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
19–21	0.232	0.453	0.567	0.619	0.699	0.750	0.259
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

(*Continued on next page*)

evidence that the differences in investment magnify the differences in initial human capital endowments, expressed as present value of lifetime wages. This is indicated by the fact that the *initial* residual levels associated with higher wage growth are sufficiently

propensity to invest is related not only to ability but also to taste parameters, such as discounting and risk aversion.

Table 3
(*Continued*)

(iii): Residuals of Mincer's wage function with fixed effects

Experience	1–3	4–6	7–9	10–12	13–15	16–18	19–21
1–3	0.141						
4–6	0.317	0.066					
	(0.000)						
7–9	−0.094	0.157	0.047				
	(0.027)	(0.000)					
10–12	−0.280	−0.209	0.218	0.047			
	(0.000)	(0.000)	(0.000)				
13–15	−0.429	−0.419	−0.267	0.072	0.056		
	(0.000)	(0.000)	(0.000)	(0.089)			
16–18	−0.481	−0.465	−0.351	−0.198	0.203	0.080	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
19–21	−0.448	−0.437	−0.351	−0.220	0.059	0.291	0.095
	(0.000)	(0.000)	(0.000)	(0.000)	(0.165)	(0.000)	

Note: Significance level in parentheses.

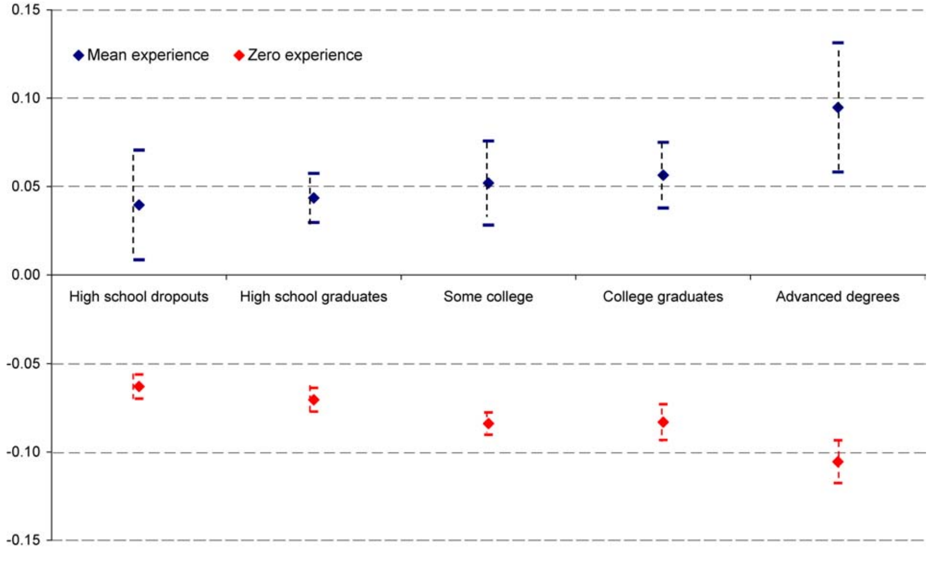

Figure 9a. Regression coefficients of individual growth rates on individual level effects, evaluated at zero experience and the individual's mean experience by education.

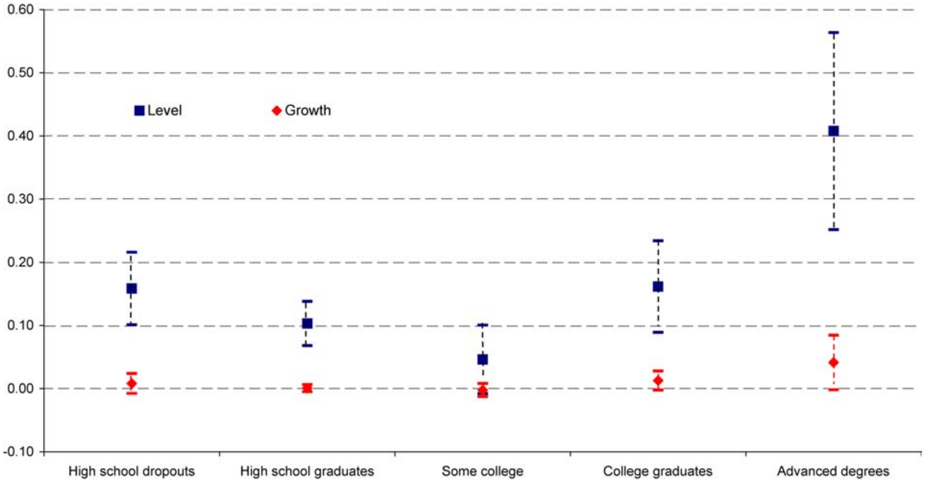

Figure 9b. Regression coefficients of individual growth effects and level effects (evaluated at the mean) on AFQT scores by education.

negative to render the total impact on the present value of lifetime earnings to be rather small.[28]

Although the investment interpretation is consistent with important features of the data on wage *levels*, it cannot explain some important feature of wage *changes*. In particular, it was noted by MaCurdy (1982) and Abowd and Card (1989) that, after accounting for the common wage growth, the growth rates of individual wages are not correlated for periods that are more than few years apart. This finding, confirmed by subsequent studies [Lillard and Reville (1999), Meghir and Pistaferri (2001), Alvarez, Browning and Ejrnaes (2001)], is also shown in Table 4a. Moreover, the correlations between short subsequent periods (one or two years) are negative. This correlation pattern is consistent with search where shocks are random, with those experiencing positive shocks less likely to exhibit high wage growth in subsequent periods. Clearly, measurement errors is another source for a negative short run correlation in individual wage growth rates. However, for sufficiently long periods (4 years) that are distant from each other one obtains a positive and significant correlation (see Table 4b) that is consistent with fixed individual growth rates, indicating that those who have above-average wage growth early in life also have above-average wage growth late in life.[29] Generally, in-

[28] Rubinstein and Tsiddon (2004), who use parents' education as a proxy for ability show that, within education groups, workers with more educated parents have higher wage levels and higher wage growth. Huggett, Ventura and Yaron (2002) show how a positive correlation between learning ability and earning ability can explain the rising variance and skewness of the earnings distribution within cohorts.

[29] The individual growth rates are estimated within cell using the regression:

$$\Delta w_{it} = b_0 + b_1 u_t + b_2 \Delta u_t + \sum d_j E_j + \theta_i + \varepsilon_{it},$$

Table 4
Variances and correlations of the residuals of the first differences of log hourly wages of full-time workers at different stages of the life cycle. NLSY, 1979–2002

a: Three-year averages

Experience (potential)	1–3	4–6	7–9	10–12	13–15	16–18
1–3	**0.092**					
4–6	−0.236	**0.077**				
	(0.000)					
7–9	−0.054	−0.228	**0.077**			
	(0.148)	(0.000)				
10–12	0.024	0.030	0.049	**0.077**		
	(0.534)	(0.266)	(0.067)			
13–15	−0.038	0.037	0.031	−0.230	**0.059**	
	(0.364)	(0.213)	(0.291)	(0.000)		
16–18	−0.058	0.073	0.038	−0.054	−0.243	**0.037**
	(0.227)	(0.032)	(0.250)	(0.067)	(0.000)	

b: Four-year averages (excluding overlapping periods)

Experience (potential)	1 to 4	6 to 9	11 to 14	16 to 19
1 to 4	**0.085**			
6 to 9	−0.076	**0.077**		
	(0.011)			
11 to 14	0.025	0.073	**0.059**	
	(0.439)	(0.004)		
16 to 19	0.067	0.016	−0.226	**0.036**
	(0.055)	(0.572)	(0.000)	

Notes: We calculate individuals' mean residuals for each cell from within cell regressions of the change in log hourly wages on experience and national unemployment rates.
Significance level in parentheses.

vestment is indicated by a positive correlation between early and late earnings, whereas search and learning imply short-term persistence with positive drift and negative correlation in wage growth. The reduction in the variance with experience is consistent with the theoretical prediction that all individuals reduce their investment to zero towards the end of life [see Lillard and Reville (1999)]. However, the NLSY sample up to 2002 maybe too young to recover end of life effects.[30]

where $\Delta w_{it} = \log(w_{it}) - \log(w_{it-1})$ or $((\log(w_{it}) - \log(w_{it-1}))/2$ is wages are reported biannually) and u_t is the national civilian unemployment rate in year t. E_j are dummies for year of potential experience. Tables 4a and 4b report the correlations between the within cells estimates of θ_i.

[30] Indeed Abowd and Card (1989), and Baker (1997), who use a wider age range, find that the variance in earnings rises at old age suggesting that individual wage shocks dominate at this stage of the life cycle.

4.3. Labor mobility and wage growth

Search theory not only competes with the theory of human capital, it also complements that theory. The challenge is to understand the interactions between these two processes. Mincer and Jovanovic (1981) provide the first attempt to integrate these processes. They describe the potential impact of search as follows "Perhaps the best way to summarize the life cycle relation between wages and mobility is to recognize that initial (first decade ?) job search has two major purposes: to gain experience, wages, and skills by moving across firms; and to find sooner or later a suitable job in which one can settle and grow for along time. The life cycle decline in mobility is, in part, evidence of successful initial mobility, an interpretation which is corroborated by corresponding life cycle growth in wages" [Mincer and Jovanovic (1981, p. 42)].

To identify the actual impacts of search and investment, they consider two different aspects of work experience, tenure in a given firm, T, and general work experience, X. They then examine two jointly determined outcomes; the wage, $w(T, X)$, and the separation rate, $s(T, x)$. The latent variables in this system are investments in general and firm-specific training and search. They use the NLS panel data and run regressions of wages and separations on tenure in the current job and potential work experience. To partially correct for the endogeneity of tenure, they add the number of past moves across firms as an indicator of individual "propensity to move".

Their main results are:

- Tenure has a separate positive and declining effect on wages, which is as important as the effect of total work experience. Tenure effects are much more important for young workers.
- Experience and tenure have negative impacts on separation, but the negative effect of tenure is much larger.
- Past moves have positive effects on separation, suggesting heterogeneity, but have only weak negative effects on wages.
- Controlling for both experience and tenure, education has a negative effect on mobility.
- The positive impact of schooling on wages is unaffected by the inclusion of mobility variables such as tenure and past moves, but the experience effects among young men are reduced substantially. This suggests that search mainly affects the size and interpretation of the experience effect but has little bearing on returns from schooling.

Subsequent work in this area tried to address the potential biases that arise when estimating the tenure effect and the impact of occupational moves. Potential biases arise from a variety of selection issues (i.e., in what ways are stayers different from movers) and in part from the assumed imperfect information and specific investments that create relational rents and give scope to bargaining and other noncompetitive behavior. A rather broad range of estimates for the size of the tenure effects have been obtained, ranging from approximately 7 to 35 percent per ten years of seniority [see Topel (1991), Altonji and Williams (1998, 2004), Dustmann and Meghir (2005)]. Data on wage loss following plan closure also indicate that the loss of wages is higher for workers with more tenure, yielding a tenure effect of about 14 percent [see Farber (1999)]. A positive tenure effect is often attributed to firm-specific human capital that is shared if the worker stays with the firm and lost if he changes employers, although it is not entirely clear why and how wage growth should respond to the accumulation of such specific capital.

A simple indication of the complexity of the relationship between wage growth and mobility is that, on the average, wage growth is associated with mobility, yet when we look at individual data, movers exhibit lower wage growth than stayers (see Figure 10). There are several possible explanations for this discrepancy: (1) If moving is a personal attribute, then firms are less likely to invest in prospective movers. (2) If jobs differ by the quality of match, successful and more productive matches are less likely to come apart. (3) If the firm is subject to exogenous shocks, the better workers are selected to stay with the firm. (4) If the continuation of the match is jointly profitable, the sharing of the gains will depend on outside options. Therefore, the *threat* of mobility rather than realized mobility can cause wage growth; much of the benefit of this threat is captured by the stayers. This threat is reflected by the *average* trends in mobility within a cohort.

Topel and Ward (1992), who examined the mobility and wage growth of young workers, find that

- Wage growth within firms is quite high (7 percent on average) and declines with both tenure and experience.
- Jobs that are going to last longer currently offer higher wage growth.
- Wage growth across jobs is substantial (20 percent on the average) and declines with tenure (at previous job) and experience.

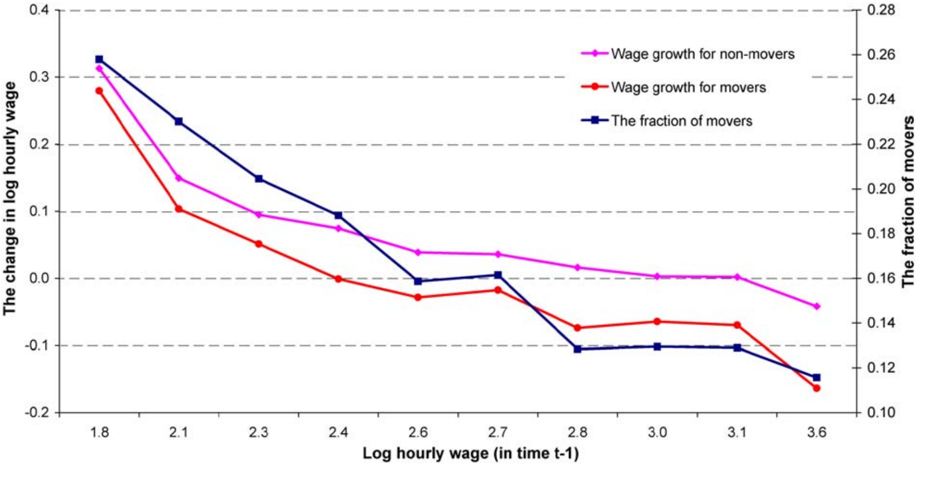

Figure 10. Fraction of movers and annual growth rates in hourly wages of movers and stayers by hourly wage in the previous year, NLSY, 1979–2000.

- Higher wage growth upon transition is obtained when one moves to a job with longer prospective tenure.
- The exit rate from a given job declines with experience and the wage level. However, conditional on the wage, the effect of experience on the job exit rate is positive.

Together these findings provide strong support for the importance of search at early stages of the worker's career.

Changes in occupation and industry are also channels for wage growth. If one ranks occupations or industries by their average wage level at the "prime" ages, 36–45, then we can identify the direction of moves on this scale. We find that the occupation and industry changes of less-educated workers involve transitions to higher paying occupations, while highly educated workers move across similar occupations and industries in terms of their mean wage.[31] In this respect, there is substitution between learning in school and on the job (see Figures 11a and 11b). In contrast, highly educated workers obtain higher wage growth when they change employers, suggesting that education and search are complements.[32] These results are consistent with the findings of Sicherman

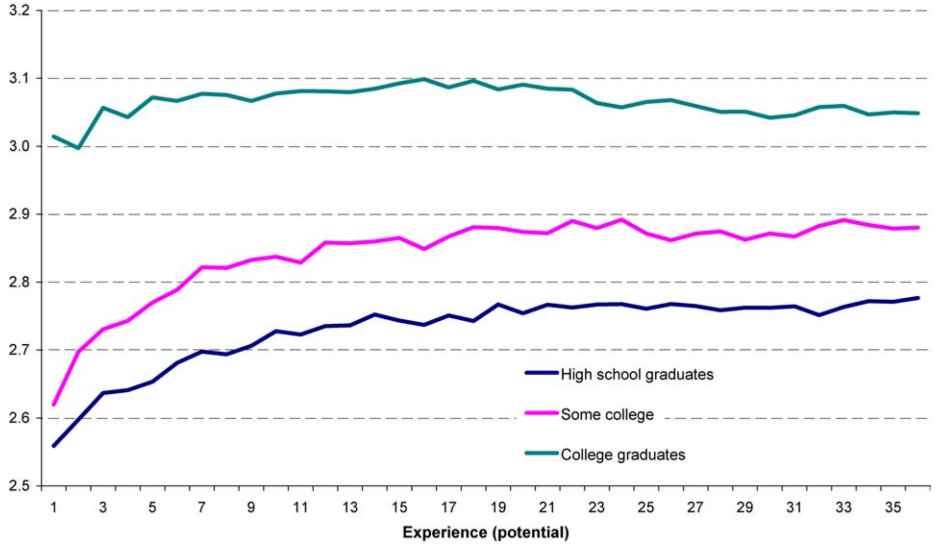

Figure 11a. Mean hourly wages (in logs) of prime aged workers (36–45) in the currently held industry and occupation, by education and experience, CPS-ORG, 1998–2002. Occupation.

[31] To examine moves across industry and occupation, we use the CPS monthly files from January 1998 to December 2002. Overall, we have in our data 473 occupation categories and 236 industries.

[32] Holding constant experience and previous wage, movers in the NLSY with higher than college degree have the same wage growth as comparable stayers. Movers, with lower levels of schooling have a substantially lower wage growth than comparable stayers.

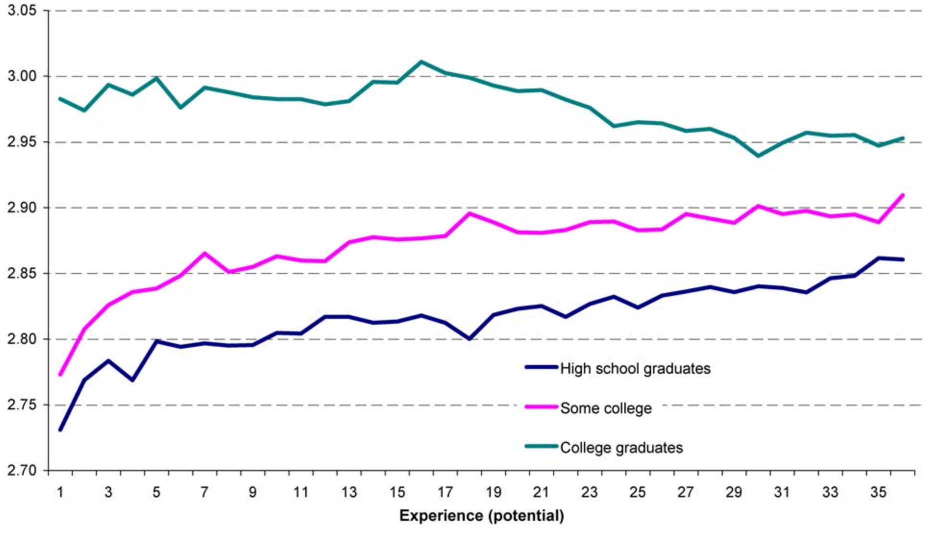

Figure 11b. Mean hourly wages (in logs) of prime aged workers (36–45) in the currently held industry and occupation, by education and experience, CPS-ORG, 1998–2002. Industry.

(1991) and Neal (1995, 1999) that educated workers are less likely to make a career change and that they also experiment with fewer employers prior to making such a change. A partial explanation is that educated workers learn about their ability in school, which facilitates their career choice. However, educated workers may take more time to find an employer that matches their skills. In fact, workers that report that their education exceeds the requirements of the job they hold are, on average, more educated and less experienced.

One must bear in mind that wage gains or losses that one observes upon job change are partial and possibly misleading indicators of the total value of such moves because workers may anticipate consequences that occur later in their career. Studies of mobility patterns over the business cycle show that movers who obtained wage gains during booms often leave their new jobs and suffer a wage loss during recession [see Keane, Moffitt and Runkle (1988), Barlevy (2001)]. There is, however, no evidence that young movers accept jobs in low-wage industries in exchange for future prospects in those industries [see Bils and McLaughlin (2001)].[33]

[33] Rubinstein and Tsiddon (2003) show that the effect of recessions on labor market outcomes varies by education and parents' education. While educated workers who were born to better-educated parents do not lose wages or jobs during recessions, less educated workers lose both.

4.4. Learning

When employers and workers are uncertain about each other's attributes, it takes time to reduce this uncertainty through experimentation. Such learning can occur within a firm or in the market at large.

As noted by Jovanovic (1979b), learning at the firm level can be inferred from the shape of the hazard function of leaving the firm. That is, if workers and firms learn about the quality of the match after they have spent an initial period together, then the weak matches terminate and the good ones survive. As time passes, learning has been accomplished and the proportion of good matches rises, so that the hazard function first rises and then declines. This is a rather sharp test because a sorting model based on the survival of the fittest usually implies a declining hazard. The hazard function in Figure 12 displays such a pattern, showing that the probability of separation conditional on length of employment peaks at about 15 months. A similar finding is reported by Booth, Francesconi and Garcia-Serrano (1999). In contrast, the data on young men used by Topel and Ward (1992) show a decline in the hazard by tenure (and experience) right from the beginning of the employment relationship. This, of course, does not exclude experimentation but shows that sorting is more important.[34]

As noted by Farber and Gibbons (1996) and Altonji and Pierret (2001), public learning can be inferred from the impact on wages of individual attributes that are not

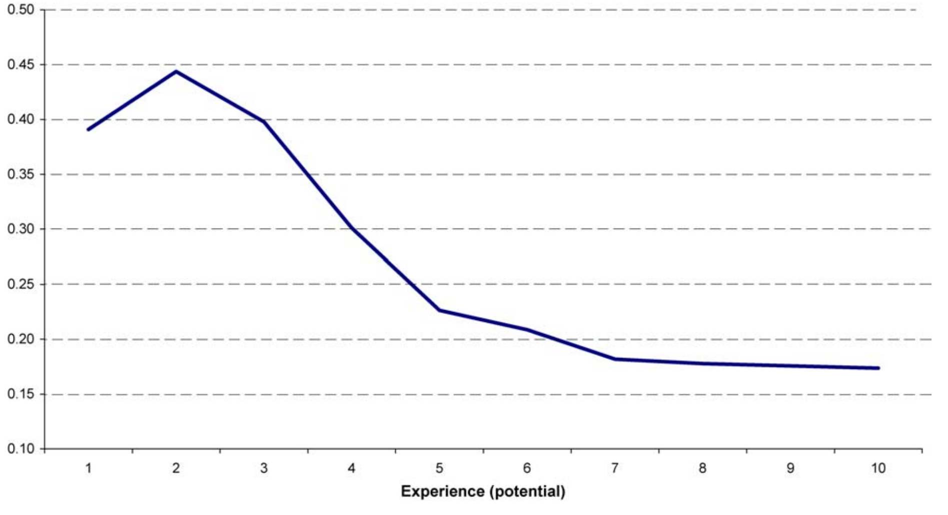

Figure 12. Hazard function of separation from current employer (in annual terms), NLSY, 1979–2000.

[34] It is interesting to note that a rising hazard function that peaks after 3 years was found in the context of divorce [see Weiss and Willis (1997)], suggesting perhaps that it is somewhat more difficult (or useful) to learn about the quality of marriage than about the quality of the job.

directly observed by employers. As time passes and employers observe the worker's performance, they learn about the worker's true productivity and the impact on wages of variables that are observed by the researcher but not by the firm (such as AFQT) increases, while the impact on wages of early signals of ability (such as schooling) declines. In Figures 13a to 13d, we show the marginal impact of AFQT on earning by experience within education groups.[35] The graphs show an increase in the impact of AFQT at early years of experience, especially for high school graduates, suggesting that learning about ability is more relevant for this group.[36] A further indicator of interest is race or ethnicity, which employers may use as a predictor of ability. In Table 5 we show that the increase in the impact of AFQT and the decline in the effect of schooling over the life cycle are substantially higher for blacks and Hispanics. This suggests initial racial statistical discrimination which gradually dissipates, as employers learn about individual ability.

Generally speaking, it is relatively difficult to tease the impact of learning from the data based on the impact of AFQT scores on wage growth.[37] Apart from problems of

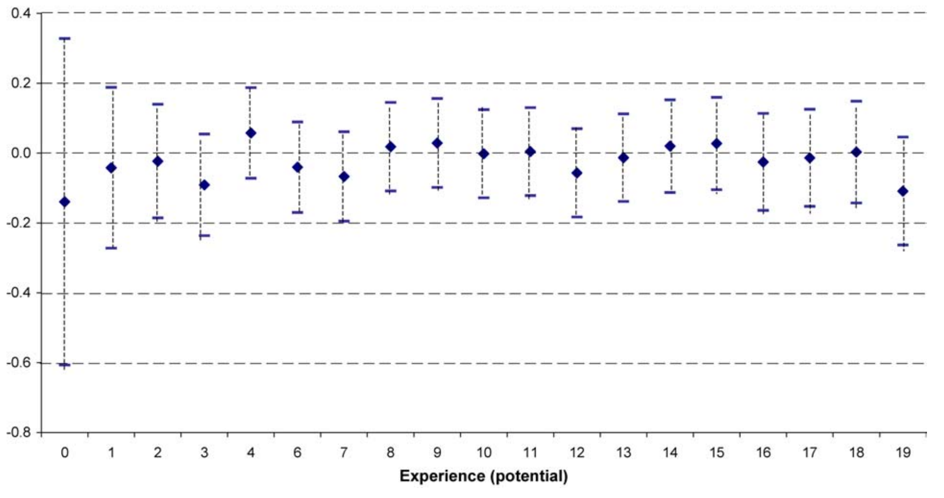

Figure 13a. The effect of AFQT on log hourly wage, by experience and education, point estimates and confidence intervals (relative to the AFQT effect at 5 years of experience), NLSY, 1979–2000. High school dropouts.

[35] Workers are classified by their completed schooling as of age 30. In each education group, and for each year of experience, we run regression with AFQT scores and year effects as explanatory variables. The figures record the estimated coefficient on the AFQT score.

[36] Lange (2003) also finds that employers' learning is concentrated at the early part of the worker's life cycle.

[37] Farber and Gibbons (1996) and Altonji and Pierret (2001) get sharper results by using more heterogeneous samples that include women and blacks, restricting the coefficients of AFQT* experience and schooling* experience to be common across groups.

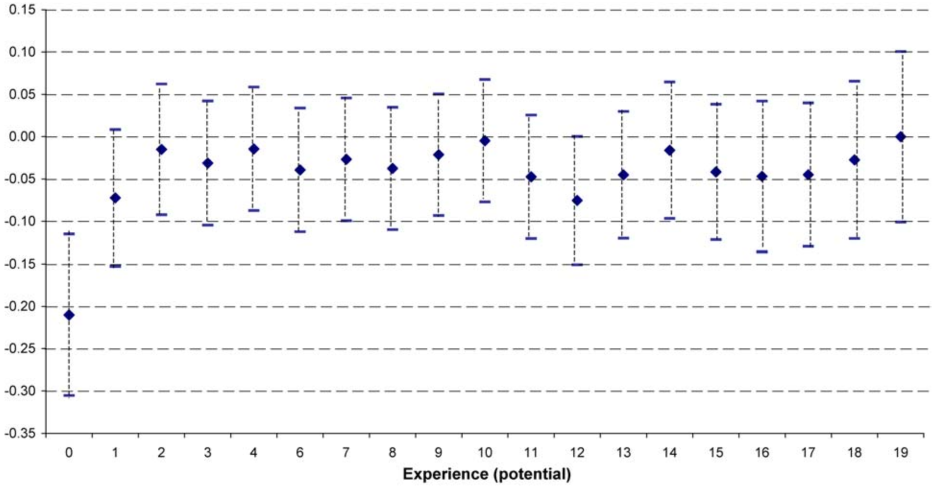

Figure 13b. The effect of AFQT on log hourly wage, by experience and education, point estimates and confidence intervals (relative to the AFQT effect at 5 years of experience), NLSY, 1979–2000. High school graduates.

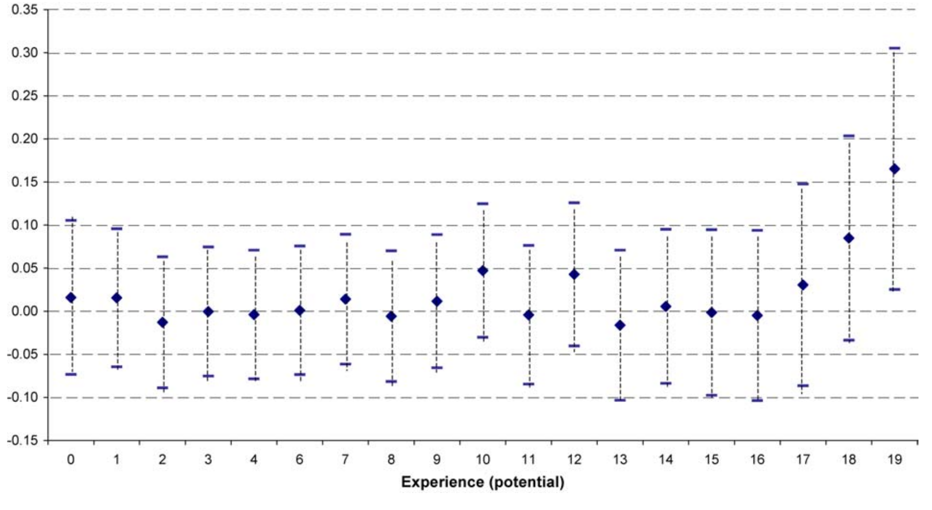

Figure 13c. The effect of AFQT on log hourly wage, by experience and education, point estimates and confidence intervals (relative to the AFQT effect at 5 years of experience), NLSY, 1979–2000. Some college.

separating learning from investment, where AFQT as an indicator of ability can affect both level and growth of wages, there are some deeper problems related to the connections between indicators of ability, such as AFQT, and wages. Willis and Rosen

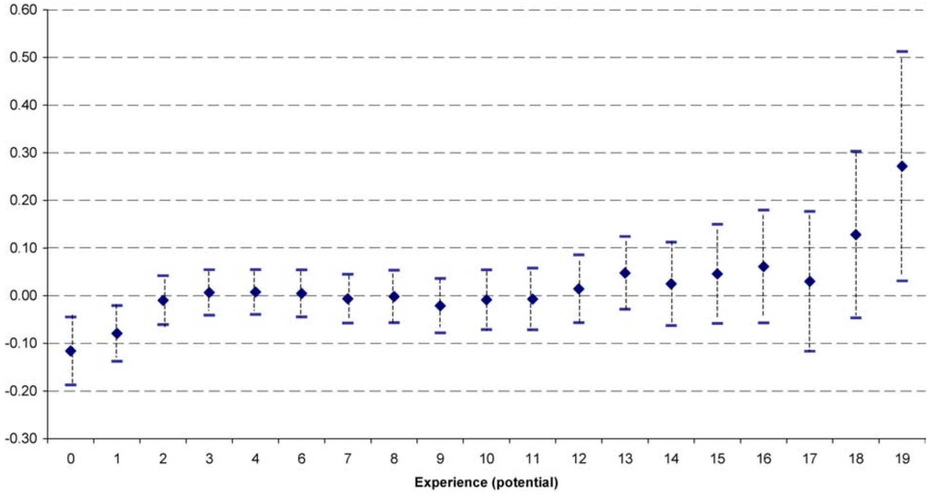

Figure 13d. The effect of AFQT on log hourly wage, by experience and education, point estimates and confidence intervals (relative to the AFQT effect at 5 years of experience), NLSY, 1979–2000. College graduates and advanced degrees.

(1979), Heckman and Rubinstein (2001) and Heckman, Hsee and Rubinstein (2003) have shown that a two factor model that recognizes the role of comparative advantage is more suitable for explaining schooling choices and wage outcomes. Figures 9a and 9b show the strong positive interaction between schooling and AFQT, which suggests that ability is more important among workers who are more educated and thus placed at more "responsible" jobs. Alternatively, the interaction indicates that high-ability individuals who do not acquire high levels of schooling may be lacking valuable non-cognitive traits. Similar issues arise in the context of the impact of AFQT on wage growth. It is quite possible that, conditional on a low level of schooling, high AFQT indicates that the worker is lacking in some other important dimension, such as motivation; as time passes this is confirmed by performance. This substitution may explain the low impact of AFQT among workers with some college and the initially negative interaction between AFQT and experience for this group.

Learning can also influence the variance of wages within a cohort of workers, as workers are gradually sorted out. It is generally difficult to separate this force for increasing variability from other considerations, such as investment, discussed above. In special cases, however, such a separation is possible. An interesting example is when workers move to a new labor market and can be followed based on their time spent in the new country. Eckstein and Weiss (2004) provide such an analysis for the wave of immigration from the former USSR to Israel during 1990–2000. The issue in this case was that employers were uncertain about the quality of schooling received in the former USSR, a factor that affects all immigrants, as well as the quality of particular immigrants. The results show that initially, all immigrants are treated alike and receive the

Table 5
Mincer's wage equation with AFQT, race and ethnicity males NLSY, 1979–2000

Variables	OLS			Fixed effects		
	All	Whites	Blacks and Hispanics	All	Whites	Blacks and Hispanics
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Black	**−0.093**	–	**−0.103**	–	–	–
	(0.020)		(0.028)			
Hispanic	**0.005**	–	–	–	–	–
	(0.023)					
AFQT	**0.043**	**0.083**	**0.043**	–	–	–
	(0.014)	(0.019)	(0.024)			
School years completed	**0.096**	**0.082**	**0.109**	–	–	–
	(0.008)	(0.010)	(0.012)			
Experience	**0.106**	**0.089**	**0.122**	**0.098**	**0.078**	**0.119**
	(0.011)	(0.014)	(0.017)	(0.006)	(0.007)	(0.010)
Experience square[1]	**−0.024**	**−0.023**	**−0.025**	**−0.027**	**−0.027**	**−0.027**
	(0.002)	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)
Interactions						
Schooling∗experience[1]	**−0.015**	**0.001**	**−0.033**	**−0.001**	**0.015**	**−0.019**
	(0.006)	(0.008)	(0.009)	(0.003)	(0.004)	(0.005)
AFQT∗experience[1]	**0.058**	**0.018**	**0.061**	**0.053**	**0.009**	**0.074**
	(0.011)	(0.016)	(0.020)	(0.006)	(0.008)	(0.010)
Observations	24801	15430	9371	24801	15430	9371
R-squared	0.319	0.318	0.272	0.265	0.306	0.201

Notes: Robust standard errors in parentheses.
[1]Coefficients and standard errors multiplied by 10.

same wage, irrespective of the experience and schooling brought from abroad. As time passes and the market learns about the immigrant's quality, the returns for imported skills rise and immigrants are gradually sorted by their observed attributes. At the same time, the residual variance reflecting unobserved attributes rises, too. The outcome is that both the mean and variance of immigrant wages rise with time spent in the new country.

One issue of interest in learning models is whether individuals move from high risk to low risk occupations or vice versa. It has been shown by Johnson (1978) and Miller (1984) that if workers are unsure about their ability to perform a job, or about the quality of the worker–job match, young workers will willingly try out jobs where success is rare, which the more-experienced have already quit after finding out that they are unsuitable. However, Jovanovic and Nyarko (1997) have shown that if what one learns from experience is how to perform the job – rather than about one's own ability or the job's quality – then the direction of mobility is reversed. Thus, the young first try the safe jobs, as long as experience is sufficiently transferable, because it is better to learn in jobs where mistakes are less costly. In Figures 14a and 14b, we show the stan-

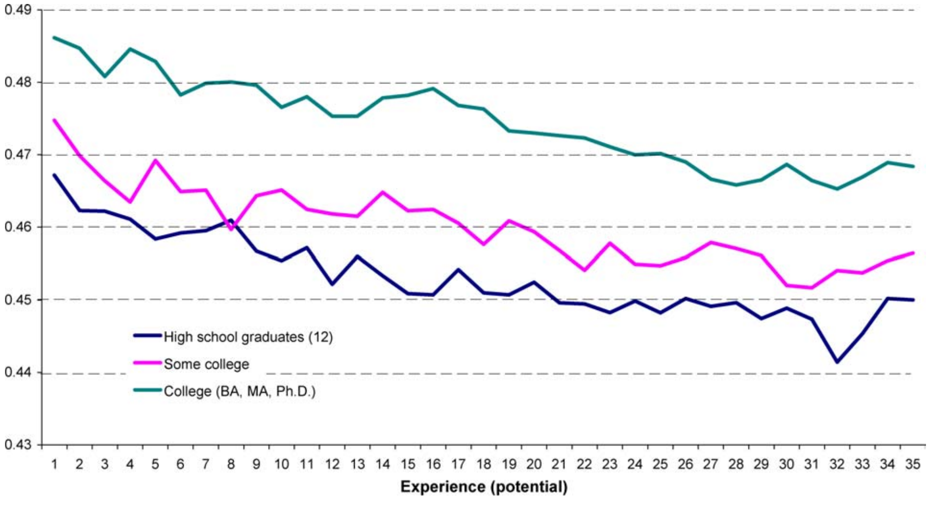

Figure 14a. Standard deviations of the log hourly wages of prime aged workers (36–45) at the industry and at the occupation in which the worker is currently employed, by education and experience, CPS-ORG, 1998–2002. Industry.

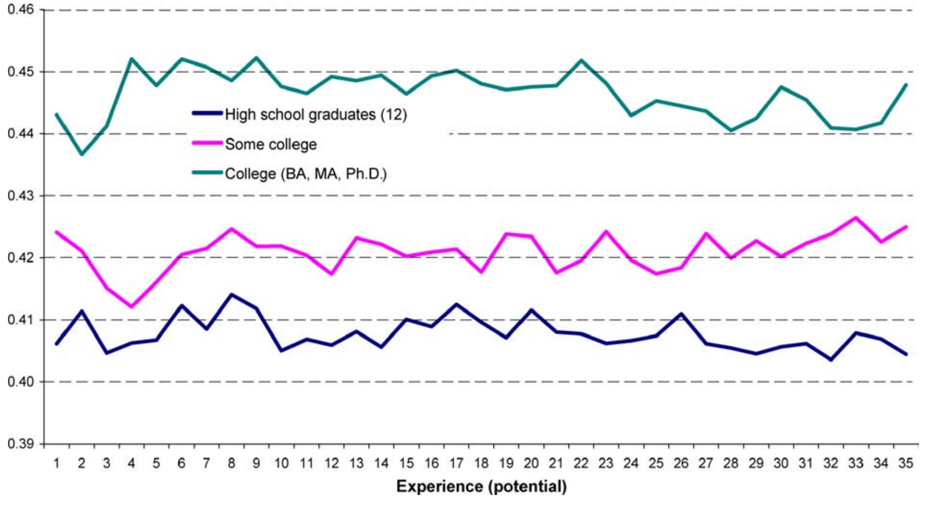

Figure 14b. Standard deviations of the log hourly wages of prime aged workers (36–45) at the industry and at the occupation in which the worker is currently employed, by education and experience, CPS-ORG, 1998–2002. Occupation.

dard deviations for the occupations and industries in which individuals are employed at different stages of their life cycle. We see that these measures of risk are stable under occupational moves but *decline* as workers change industries. The results suggest experimentation with match quality across industries.

5. Data appendix: Data and sample-inclusion criteria

5.1. The CPS individual-level repeated cross-section data set

These data come from a series of 39 consecutive March *Current Population Surveys* (hereafter: March CPS) for the years 1964 to 2002. These data provide information on employment and wages in the preceding calendar year. Thus, the annual data – taken from the CPS demographic supplement – cover the period of 1963 to 2001. The individual-level repeated cross-section data set is restricted to men aged 18 to 65 with zero (0) to forty (40) years of potential experience, where potential experience is defined as age-6-school years completed.

The main advantage of the March CPS is that micro data samples are available from the mid-1960s onward. On the minus side, the March CPS has no "point-in-time" measure of the wage rate. Wage rates, in many of studies using the March CPS data, are often constructed by dividing total annual earnings in the previous year by an estimate of weeks or hours of work. The task is made more difficult by the absence of information on usual hours of work per week prior to 1976. For these reasons we further restrict this sample to include – Full-Time-Full-Year workers (hereafter: FTFY) – full-time workers (35+ hours per week) who reported working at least 51 weeks of the previous year.

The wage measure in the March CPS data set that we use throughout this paper is the average weekly wage computed as total annual earnings divided by total weeks worked. Top coding has been changed over the years. Until the 1995 survey, the imputed wages/earnings of top-coded workers were set to equal the cutoff point. Since 1996, the imputed wages for the top-coded group are based on the conditional mean earnings of these workers conditional on characteristics such as race, gender and region of residence. In order to deal with the top-coding issue, we employ a unified rule for all years. We calculate for each worker his rank/position on the wage distribution for the year observed and exclude those belonging to either the lower 2 percent or the top 2 percent each year.

Observations are divided by completed schooling, when interviewed, into five categories: (i) high school dropouts – less than twelve grades, (ii) high school graduates (iii) some college completed, (iv) college graduates with 16 years of schooling (BA) and (v) college graduates with advanced/professional education (MBA, PhD).

5.2. The CPS monthly longitudinally matched data

The vast majority of empirical analyses of the Current Population Surveys use either a single cross-section data point or a series of consecutive CPS surveys, treating the latter as a series of repeated cross-sections. The CPS data have, in fact, a longitudinal component. In this paper we take advantage of the CPS basic monthly files – a probability sample of housing units in the US – to construct a panel data set.

The CPS divides housing units into 8 representative sub-samples called "rotation groups". Each unit is interviewed for 4 consecutive months, followed by a break lasting two quarters, and again for another four monthly interviews. Overall, each unit is interviewed for 8 times over 16 months. The CPS monthly files we employ – from the years 1998 to 2002 – include a set of identifier variables that enables us to follow the same housing unit over 16 months. If there is no change in the composition of individuals residing in a particular unit, we have a panel of individuals. Yet, since people do switch locations, it might be the case that the same id number was shared by 2 (or more) individuals over time. Therefore, we follow the Madrian and Lefgren (1999) procedure, whereby individuals are identified in our panel data not only by their id number but also by matching a set of time-invariant characteristics. This procedure make us quite confident that we do not combine different persons into one artificial observation.

Data on schooling, employment, occupation and industry, are available for all interviews. However weekly wage data is collected only during the fourth and the eighth interview – among what is known as the "outgoing rotation groups" hereafter (ORG). We construct two samples. The main sample includes only workers participating in all interviews. This sample is used for the analysis of transitions between industries and occupations. Our second sample is taken from the ORG sample restricted to full-time workers, not enrolled in school and with two wage data points. We exclude observations with a reported hourly wage lower than \$4 or higher than \$2000 (adjusted for 2000 CPI). This sample is used to study wage growth of individuals.

5.3. The panel study of income dynamics

The Panel Study of Income Dynamics (PSID) is a longitudinal, nationwide survey of a representative sample of individuals and the families in which they reside. The PSID began in 1968 with approximately 4,800 white and black households and approximately 18,000 individuals. The sample had expanded as original members formed additional families over the years. We restrict our sample to US born white males aged 21 to 65 during the work year, with non missing demographics. When we discuss wage data, we exclude workers with a reported hourly wage lower than \$4 or higher than \$2000 (adjusted for 2000 CPI) and individuals who worked less than 35 weeks or less than 1000 annual hours. When using wage differences, we restrict the sample according to these cretiria in both consecutive years. Observations are divided by completed schooling into five categories – similar to our definitions using the CPS data.

6. National Longitudinal Survey of Youth (NLSY)

The micro data we use are from the 1979–2000 waves of the National Longitudinal Survey of Youth (NLSY).[38] The NLSY includes a randomly chosen sample of US youths and a supplemental sample that includes Black, Hispanic, and non-Black, non-Hispanic economically disadvantaged young people. Interviewees have been surveyed annually since the initial wave of the survey in 1979, when sample members all ranged between age 14 and 21 in 1979. The military sub-sample and the non-black, non-Hispanic disadvantaged samples are excluded. We further exclude observations with missing data regarding own or parents' education, Armed Forces Qualification Test score (hereafter AFQT), or labor market outcomes. In order to guarantee that AFQT test scores were not influenced by school attendance, AFQT scores are gender-age-school-adjusted (standardized within birth year cohort to mean 0, variance 1). When studying labor market outcomes we exclude individuals enrolled in schooling in the given year. We group respondents into five education categories: high school dropouts, high school graduates (including GED graduates), some college (SC), college graduates and individuals with advanced degrees.

When we discuss wage data, we further exclude workers with a reported hourly wage lower than \$4 or higher than \$2000 (adjusted for 2000 CPI) and individuals who worked less than 35 weeks or less than 1000 annual hours. When using wage differences, we restrict the sample according to these criteria in both consecutive years.

References

Abowd, J., Card, D. (1989). "On the covariance structure of earnings and hours changes". Econometrica 57, 411–445.

Abraham, K., Farber, H. (1987). "Job duration, seniority and earnings". American Economic Review 77, 278–279.

Altonji, J., Pierret, C. (2001). "Employer learning and statistical discrimination". The Quarterly Journal of Economics 116, 313–350.

Altonji, J., Williams, N. (1998). "The effect of labor market experience, job seniority and wage growth". Polachek, S. (Ed.), Research in Labor Economics 17, 233–276.

Altonji, J., Williams, N. (2004). "Do wages rise with job seniority? A reassessment". Industrial and Labor Relations Review. In press.

Alvarez, J., Browning, M., Ejrnaes, M. (2001). "Modeling income processes with lots of heterogeneity". Unpublished Manuscript, Institute of Economics, Copenhagen University.

Baker, G., Gibbs, M., Holmstrom, B. (1994a). "The internal economics of the firm: Evidence from personnel data". Quarterly Journal of Economics 109, 881–919.

Baker, G., Gibbs, M., Holmstrom, B. (1994b). "The wage policy of a firm". Quarterly Journal of Economics 109, 921–956.

Baker, M. (1997). "Growth rate heterogeneity and the covariance structure of life time earnings". Journal of Labor Economics 15, 837–878.

[38] Exceptions are Tables 3 and 4 where, in order to increase the number of observations at later ages, we also use the survey year 2002.

Barlevy, G. (2001). “Why are the wages of job changers so procyclical?”. Journal of Labor Economics 19, 94–135.

Beaudry, P., Di Nardo, J. (1991). “The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data”. Journal of Political Economy 99, 665–688.

Becker, G. (1975). Human Capital. Columbia University Press.

Ben-Porath, Y. (1967). “The production of human capital and the life cycle earnings”. Journal of Political Economy 75, 352–365.

Bils, M., McLaughlin, K. (2001). “Interindustry mobility and the cyclical upgrading of labor”. Journal of Labor Economics 19, 94–135.

Blau, F., Kahn, L. (1981). “Causes and consequences of layoffs”. Economic Inquiry 19, 270–296.

Blundell, B., MaCurdy, T. (1999). “Labor supply: A review of alternative approaches”. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3b. North-Holland, Amsterdam.

Booth, A., Francesconi, M., Garcia-Serrano, C. (1999). “Job tenure and job mobility in Britain”. Industrial and Labor Relations Review 53, 43–70.

Booth, A., Francesconi, M., Frank, J. (2002). “Temporary jobs: Stepping stones or dead ends?”. Economic Journal 112, F585–F606.

Borjas, G. (1985). “Assimilation, changes in cohort quality, and the earning of immigrants”. Journal of Labor Economics 3, 463–489.

Burdett, K. (1978). “Job search and quit rates”. American Economic Review 68, 212–220.

Card, D. (1999). “The causal effect of education on earnings”. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3a. North-Holland, Amsterdam.

Card, D. (2001). “Estimating the return to schooling: Progress on some persistent econometric problems”. Econometrica 69, 1127–1161.

Chiappori, P., Salenei, B., Valentin, J. (1999). “Early starters versus late beginners”. Journal of Political Economy 102, 731–760.

Dustmann, C., Meghir, C. (2005). “Wages, experience and seniority”. Review of Economic Studies 72, 77–108.

Eckstein, Z., Weiss, Y. (2004). “On the wage growth of immigrants: Israel 1990–2000”. Journal of the European Economic Association 2, 665–695.

Farber, H. (1999). “Mobility and stability: The dynamics of job changes in labor markets”. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3b. North-Holland, Amsterdam.

Farber, H., Gibbons, R. (1996). “Learning and wage dynamics”. Quarterly Journal of Economics 111, 1007–1048.

Felli, L., Harris, C. (1996). “Learning wage dynamics and firm-specific human capital”. Journal of Political Economy 104, 838–868.

Felli, L., Harris, C. (2003). “Firm specific training”. Unpublished Manuscript, London School of Economics.

Flinn, C. (1997). “Equilibrium wage and dismissal processes”. Journal of Business and Economic Statistics 15, 221–236.

Gibbons, R., Waldman, M. (1999a). “A theory of wage and promotion dynamics inside firms”. Quarterly Journal of Economics 114, 1321–1358.

Gibbons, R., Waldman, M. (1999b). “Careers in organizations: Theory and evidence”. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3b. North-Holland, Amsterdam.

Griliches, Z. (1977). “Estimating the returns to schooling: Some econometric problems”. Econometrica 45, 1–22.

Haider, S. (2001). “Earning instability and earning inequality of males in the United States: 1967–1991”. Journal of Labor Economics 19, 774–798.

Hall, R. (1982). “The minimum wage and job turnover in markets for young workers”. In: Freeman, R., Wise, D. (Eds.), The Youth Labor Market. University of Chicago Press, Chicago.

Harris, M., Holmstrom, B. (1982). “A theory of wage dynamics”. Review of Economic Studies 49, 315–333.

Harris, M., Weiss, Y. (1984). “Job matching with finite horizon and risk aversion”. Journal of Political Economy 92, 758–779.

Heckman, J., Honore, B. (1990). "The empirical content of the Roy model". Econometrica 58, 1121–1149.
Heckman, J., Hsee, J., Rubinstein, Y. (2003). "The GED is a mixed signal: The effect of cognitive skills and personality skills on human capital and labor market outcomes". Unpublished Manuscript, Department of Economics, University of Chicago.
Heckman, J., Lochner, L., Taber, C. (1998). "Explaining rising wage inequality: Explanations with a dynamic equilibrium model of labor earnings with heterogenous agents". Review of Economic Dynamics 1, 1–58.
Heckman, J., Lochner, L., Todd, P. (2001). "Fifty years of Mincer earnings regressions". Unpublished Manuscript, Department of Economics, University of Chicago.
Heckman, J., Rubinstein, Y. (2001). "The importance of noncognitive skills: Lessons from the GED testing program". American Economic Review 91, 145–149.
Huggett, M., Ventura, G., Yaron, A. (2002). "Human capital and earning distribution dynamics". Unpublished Manuscript, Department of Economics, Georgetown University.
Johnson, W. (1978). "A theory of job shopping". Quarterly Journal of Economics 93, 261–277.
Jovanovic, B. (1979a). "Job matching and the theory of turnover". Journal of Political Economy 87, 972–990.
Jovanovic, B. (1979b). "Firm specific capital and turnover". Journal of Political Economy 87, 1246–1260.
Jovanovic, B. (1984). "Matching, turnover and unemployment". Journal of Political Economy 92, 108–122.
Jovanovic, B., Nyarko, Y. (1997). "Stepping Stone mobility". Carnegie-Rochester Conference on Public Economics 46, 289–325.
Katz, L., Autor, D. (1999). "Changes in the wage structure and wage inequality". In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3a. North-Holland, Amsterdam.
Keane, M., Moffitt, R., Runkle, D. (1988). "Real wages over the business cycle: Estimating the impact of heterogeneity with micro data". Journal of Political Economy 96, 1232–1266.
Lange, F. (2003). "The returns to schooling and ability during the early career: Evidence on job market signaling, employer learning and post-schooling investments". Unpublished Manuscript, Department of Economics, University of Chicago.
Lillard, L., Reville, R. (1999). "Life cycle human capital investment: New evidence on an old issue". Unpublished Manuscript, Rand Corporation.
Lillard, L., Weiss, Y. (1979). "Components of variation in panel earnings data: American scientists 1960–1970". Econometrica 47, 437–454.
MaCurdy, T. (1982). "The use of time series processes to model the error structure of earnings in longitudinal data analysis". Journal of Econometrics 18, 83–114.
Madrian, B.C., Lefgren, L. (1999). "A note on longitudinally matching current population survey (CPS) respondents". NBER Technical Working Paper 247.
Malcomson, J. (1997). "Contracts, hold-up and labor markets". Journal of Economic Literature 35, 1916–1957.
Malcomson, J. (1999). "Individual employment contracts". In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3b. North-Holland, Amsterdam.
McLaughlin, J. (1994). "Rigid wages?". Journal of Monetary Economics 34, 383–414.
Meghir, C., Pistaferri, L. (2001). "Income variance dynamics and heterogeneity". Institute for Fiscal Studies Discussion Paper 01/07.
Miller, R. (1984). "Job matching and occupational choice". Journal of Political Economy 92, 1086–1120.
Mincer, J. (1958). "Investment in human capital and the personal income distribution". Journal of Political Economy 66, 281–302.
Mincer, J. (1974). Schooling, Experience and Earnings. Columbia University Press, New York.
Mincer, J. (1993). Studies in Human Capital: Collected Essays of Jacob Mincer, vol. 1. Edward Elgar.
Mincer, J., Jovanovic, B. (1981). "Labor mobility and wages". In: Rosen, S. (Ed.), Studies in Labor Markets. University of Chicago Press, Chicago.
Mortensen, D. (1986). "Job search and labor market analysis". In: Ashenfelter, O., Layard, R. (Eds.), Handbook of Labor Economics. North-Holland, Amsterdam.
Mortensen, D. (1988). "Wages separations and job tenure". Journal of Labor Economics 4, 572–586.
Mortensen, D., Pissarides, C. (1999). "New developments in models of search in the labor market". In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3b. North-Holland, Amsterdam.

Munasinghe, L. (2003). "A theory of wage and turnover dynamics". Unpublished Manuscript, Department of Economics, Barnard College.
Murphy, K., Welch, F. (1990). "Empirical age earnings wage profiles". Journal of Labor Economics 8, 202–229.
Murphy, K., Welch, F. (1992). "The structure of wages". Quarterly Journal of Economics 107, 285–326.
Neal, D. (1995). "Industry specific human capital: Evidence from displaced workers". Journal of Labor Economics 13, 653–677.
Neal, D. (1999). "The complexity of job mobility among young men". Journal of Labor Economics 17, 237–261.
Polachek, S. (2003). "Mincer's overtaking point and the lifecycle earning distribution". Unpublished Manuscript, Department of Economics, SUNY at Binghamton.
Prendergast, C. (1999). "The provision of incentives in firms". Journal of Economic Literature 37, 7–63.
Rosen, S. (1972). "Learning and experience in the labor market". Journal of Human Resources 7, 326–342.
Rosen, S. (1977). "Human capital: A survey of empirical research". In: Ehrenberg, R. (Ed.), Research in Labor Economics, vol. 1. JAI Press, Greenwich.
Rubinstein, Y., Tsiddon, D. (2003). "Born to be unemployed: Unemployment and wages over the business cycle". Unpublished Manuscript, School of Economics, Tel-Aviv University.
Rubinstein, Y., Tsiddon, D. (2004). "Coping with technological progress: The role of ability in making inequality so persistent". Journal of Economic Growth. In press.
Shaw, K. (1996). "An empirical analysis of risk aversion and income growth". Journal of Labor Economics 14, 626–653.
Sicherman, N. (1991). "Over education in the labor market". Journal of Labor Economics 9, 101–122.
Topel, R. (1991). "Specific capital, mobility, and wages: Wages rise with job seniority". Journal of Political Economy 99, 145–176.
Topel, R., Ward, M. (1992). "Job mobility and careers of young men". Quarterly Journal of Economics 107, 339–479.
Weinberg, B. (2003). "Experience and technology adoption". Unpublished Manuscript, Department of Economics, Ohio State University.
Weiss, Y. (1971). "Learning by doing and occupational specialization". Journal of Economic Theory 3, 189–199.
Weiss, Y. (1984). "Wage contracts when output grows stochastically: The role of mobility costs and capital market imperfections". Journal of Labor Economics 2, 155–174.
Weiss, Y. (1986). "The determination of life-time earnings: A survey". In: Ashenfelter, O., Layard, R. (Eds.), Handbook of Labor Economics. North-Holland, Amsterdam.
Weiss, Y., Lillard, L. (1978). "Experience, vintage, and time effects in the growth of earnings: American scientists 1960–1970". Journal of Political Economy 86, 427–448.
Weiss, Y., Sauer, R., Gotlibovski, M. (2003). "Immigration, search and loss of skill". Journal of Labor Economics 21, 557–592.
Weiss, Y., Willis, R. (1997). "Match quality, new information, and marital dissolution". Journal of Labor Economics 15, S293–S329.
Welch, F. (1969). "Linear synthesis of skill distribution". Journal of Human Resources 4, 311–327.
Willis, R. (1986). "Wage determination". In: Ashenfelter, O., Layard, R. (Eds.), Handbook of Labor Economics, vol. 1. North-Holland, Amsterdam.
Willis, R., Rosen, S. (1979). "Education and self selection". Journal of Political Economy 87 (2), S7–S36.
Wolpin, K. (1992). "The determinants of black–white differences in early employment careers: Search, layoffs, quits and endogenous wage growth". Journal of Political Economy 100, 535–560.
Wolpin, K. (2003). "Wage equation and education policy". In: Dewaterpoint, M., Hansen, L., Turnovsky, S. (Eds.), Advances in Economics and Econometrics. Cambridge University Press, Cambridge, UK.

Proof of Raw Subject Index