WORKING PAPER #372
PRINCETON UNIVERSITY
INDUSTRIAL RELATIONS SECTION
DECEMBER, 1996

Changes in Relative Wages in the 1980s: Returns to Observed and Unobserved Skills and Black-White Wage Differentials*

Kenneth Y. Chay
University of California at Berkeley

David S. Lee
Princeton University

December 9, 1996

* We are indebted to David Card, Bo Honoré, and Larry Katz for many helpful discussions and invaluable suggestions, and to Hank Farber and Jim Powell, who provided comments on an earlier draft. We also thank Dean Hyslop, Alan Krueger, Cecilia Rouse, Orley Ashenfelter, and participants of the Princeton Labor Lunch, as well as Chinhui Juhn, and participants of the 1996 NBER Summer Institute Labor Studies Workshop for helpful suggestions. Financial support from the NSF Traineeship in the Economics of Education is gratefully acknowledged.

ABSTRACT

During the 1980s, did the sharp increase in the college-high school wage differential represent a rise in the college premium, or a growth in the payoff to unmeasured "ability" or "skill"? Can the slowdown in black-white wage convergence or the widening black-white gap among young workers witnessed during the 1980s be explained by a rise in the return to pre-labor market factors correlated with race? In this study, we show that it is possible to use across-group variation in within-group wage variances from multiple periods to identify the change in the return to unobservable skill, within a relatively unrestrictive error-components model of wages. The identification does not require full specification of the time-series properties or the functional form of the errors. Male earnings data from the CPS show that there is useful variation in within-group wage variances --- enough to estimate a growth in the return to unobservable skill of about 10 to 20 percent during the 1980s. In our analysis, these magnitudes imply that even after controlling for the effects of an increase in the payoff to unobservable skill, college-educated workers still gain substantially relative to high school-educated workers, while young black men still experience a significant wage decline relative to white men during the 1980s.

Kenneth Y. Chay
Department of Economics
University of California at Berkeley
Berkeley, CA
(510) 643-8596
kenchay@econ.berkeley.edu

David S. Lee
Department of Economics
Princeton University
Princeton, NJ 08544
(609) 258-4040
davidlee@princeton.edu

1 Introduction

During the 1980s, wage inequality grew along several dimensions in the United States. Most notably, after experiencing a decline in the previous decade, the measured college-high school wage differential increased substantially during the 1980s. In addition, wage inequality within narrowly-defined demographic groups based on education and experience also rose, continuing a trend that began in the early 1970s.¹ At least two hypotheses for these developments have received wide attention in the empirical literature on the wage structure in the U.S. The first is the conjecture that rising within-group wage inequality reflects an increase in the return to unobservable “skill”.² The second is the presumption that the rapid growth in the measured college-high school wage differential represented a sharp increase in the economic return to college.

However, if there is a distinct and nonstable economic return to unobservable skill, the likelihood of a correlation between unobservable skill and education introduces serious problems in the identification of true changes in the college premium. In principle, movements in the college-high school wage differential could provide a misleading picture of changes in the returns to a college education. More specifically, an increase in measured educational wage differentials could potentially be driven by changes in the return to unmeasured productivity components correlated with, but not the result of, educational attainment (e.g., innate ability, family background influences).

¹ Major changes in the U.S. wage structure have been documented in Blackburn, Bloom, and Freeman (1990), Murphy and Welch (1992), Katz and Murphy (1992), and Juhn, Murphy, and Pierce (1993).

² See the survey by Levy and Murnane (1992) for other explanations that have been offered for rising within-group wage dispersion.

When this possibility is allowed, identification of *both* 1) the extent of the “unobserved skill” (or omitted-ability) bias at a given point in time *and* 2) the growth of the payoff to unobserved skill are necessary to identify the true change in the college premium, even in a simple model which allows for only one component of unobserved skill. Without information on these two magnitudes, the true change in the college premium is indeterminate. However, identification of the change in the return to unobservable skill allows one to generate a range of plausible estimates of the change in the return to college corresponding to varying beliefs about the unobserved skill bias.

In this study, we show that it is possible to use across-group variation in within-group wage variances from multiple periods to identify the change in the return to unobservable skill, within a relatively unrestrictive error-components model of the wage-generating process. The identification strategy is sufficiently unrestrictive that it accommodates a nonzero correlation between observable and unobservable components of skill, and avoids specification of the time-series properties of every error component in the wage process. This framework allows us to empirically assess the implications of growing wage dispersion on conventional estimates of the changes in the college premium during the 1980s. It also allows us to empirically evaluate the hypothesis that real reductions in race-based labor market discrimination were masked by a changing return to unobservable skill during the 1980s.

Earnings data for males from the Current Population Survey show that there is useful variation in within-group wage variances across narrowly-defined demographic groups. This variation across groups and over time allows us to estimate the growth in the return to unobserved skill, which we estimate to be a 10 to 20 percent increase during the course of

the 1980s. We find that this magnitude is too small to support the hypothesis that the sharp rise in college-high school wage differentials is attributable to an increasing importance of unobserved components of productivity correlated with education. Our estimates are also too low to support the claim that the growth in the payoff to unobserved skill worked to completely offset progress in black-white wage convergence during the 1980s.

The paper is organized as follows. Section 2 presents a parsimonious econometric model of the wage-generating process which allows for a distinct return to unobservable skill. We use this framework to 1) point out the difficulties in estimating changes in the college premium or discrimination, in the presence of a return to unobserved skill, 2) characterize and quantify the potential biases of conventional estimates, and 3) illustrate our main identification strategy, which is described in detail in section 3. Section 4 discusses our choice of data used for this analysis, as well as specific features that are relevant to our application. Section 5 explains the specifics of estimation, and reports the results, and section 6 summarizes our findings.

2 (Non-)identification of the college premium

In this section, we briefly present three stylized facts about the wage structure in the 1980s. Within a simple econometric framework, we illustrate to what extent a change in the return to unmeasured components of skill could affect conventional estimates of changes in the return to college.

2.1 Relative wages in the 1980s: returns to college or returns to unobserved skill?

Figure 1 summarizes three stylized facts about the changing wage structure of the 1980s. It shows the movement in the measured college-high school wage differential, the black-

white wage differential, and the standard deviation of “within-group” wages. These are computed using the Current Population Survey (CPS) Merged Outgoing Rotation Group files from 1979 to 1991.³ The figure shows that the 1980s witnessed a dramatic increase in the college-high school wage differential, a growth of about .20 log points. It also shows that the measured black-white wage differential (controlling for main effects in education and experience) experienced a .04 log point decline, from -.15 to about -.19.⁴ Finally, the residual standard deviation of wages rose about .04 log points during the decade. If $\log(\text{wages})$ were distributed normally, this would imply that the 90-10 percentile wage differential, among workers with the same observable characteristics (education, experience, race), rose by about .10 log points.

There is little direct evidence for the cause of rising within-group wage inequality, but researchers have suggested that it may be reflecting an increase in the return to unobserved “skill” — that is, workers may have skills which have market value, but are not captured by the demographic or educational information which we obtain from surveys such as the CPS.⁵ Moreover, the “price” of these skills may be rising, a hypothesis that would be consistent with the observation of rising within-group wage dispersion.

Researchers have proposed and attempted to evaluate various explanations for the rising college-high school wage differential in the 1980s. Many of these hypotheses are common in

³ For each year, the sample consists of white and black male wage earners between the ages of 18 and 64 with real wages (computed as the edited hourly wage if it exists, and if not, the edited usual weekly earnings divided by the usual edited weekly hours, in 1991 dollars) between \$2 and \$60 an hour. The college-high school and black-white series are coefficients on the dummy variables 16+ yrs. of schooling and black, respectively, from the following regression: $\log(\text{wage})$ on a full set of single-year potential experience (age-education-6) dummies, 3 education categorical variables (<12, 13-15, 16+), and a black dummy variable. The residual dispersion series is the root mean squared error of the regression of $\log(\text{wage})$ on a fully interacted set of dummy variables for experience, years of schooling, and race.

⁴ Movements in black-white wage differentials during the 1980s have been documented in Bound and Freeman (1989 and 1992).

⁵ Examples of studies which adopt this approach are Juhn, Murphy, and Pierce (1993) and Card and Lemieux (1993).

that they attribute the change in the college-high school wage differential to shifts in the relative demand for, or supply of, college-educated labor (see Katz and Murphy, 1992).⁶ However, in much the same way the measured college-high school wage differential may not reflect the true return to college, *changes* in the measured college-high school wage differential need not represent true *changes* in the market return to college.

This idea is instrumental in a study within a different context, by Juhn, Murphy, and Pierce (1991), who offer an alternative explanation to the hypothesis that the slowdown in black-white wage convergence in the 1980s was the result of weakened federal antidiscrimination efforts during the decade.⁷ In particular, they point out that an increase in residual wage dispersion, if it reflects an economy-wide increase in the return to unobserved skill, has implications for measured changes in the black-white wage gap if this gap, at least in part, reflects unobserved productivity differences. For example, if black men have on average less unmeasured “skill” than whites due to differences in quality of schooling, then an increase in the price of unobserved skill will cause black-white wage differentials to expand for reasons unrelated to changes in discrimination.⁸

Clearly, the same logic also applies to the interpretation of measured educational wage differentials. If the economic return to unmeasured skills which are not acquired through education increases over time — a hypothesis consistent with the observed increase in within-group wage inequality — and if the level of these skills vary by educational attainment, it

⁶ Bound and Johnson (1992) also use a relative supply and demand framework in their attempt to evaluate the various hypotheses for changes in the wage structure during the 1980s.

⁷ One potential explanation for the slowdown in black-white wage convergence is that it was the result of a legislative environment increasingly opposed to affirmative action and equal employment opportunity policies. See Smith (1993).

⁸ Juhn, Murphy, and Pierce (1991) hypothesize that during the 1980s, true labor market discrimination continued to fall at the same rate as during the 1970s, but was offset by the an increase in the return to skill, which worked against blacks during the decade. See Card and Lemieux (1994) and Grogger (1996) for two recent studies that attempt to evaluate this explanation for the slowdown.

follows that the measured college-high school wage differential could increase even when the true economic reward to a college education remains constant.⁹

The next section illustrates that even in the simplest model that incorporates this possibility, the true change in the college premium is unidentified in the absence of extra information about the magnitudes of the unobserved skill bias as well as the growth in the return to unobserved skill.

2.2 Non-identifiability of changes in the return to observed skill

The purpose of this section is to outline the most parsimonious econometric model of the wage-generating process that allows for an unobserved skill bias, and a distinct and possibly changing return to unobserved skill. In addition, given the literature's interpretation of rising residual wage dispersion, the model also has the observable implication that within-group wage dispersion will rise whenever the return to unobservable skill rises, all other things being equal. Even though we seek to develop the simplest such model, we also attempt to minimize the number of arbitrary assumptions about functional form or other stochastic properties (such as the time series properties) of the various random components of the wage process.¹⁰

Suppose that we have two groups of workers, *HS* and *COL*, those with a high school education and those who completed college or higher, and that log-wages are determined by

⁹ Blackburn and Neumark (1993) and Taber (1996) both use NLSY data to examine the potential impact of unobserved ability biases on estimates of changes in the returns to education in the 1980s. Blackburn and Neumark find little evidence of omitted-ability biases when using ASVAB test scores as a proxy for ability. Taber, on the other hand, claims that the increase in the college premium during the 1980s is virtually eliminated after controlling for the rise in the payoff to unobserved ability.

¹⁰In particular, we want to avoid identifying any parameter of interest through functional form or times series properties assumptions.

the following equation:

$$w_{ijt} = r_t k_j + \psi_t a_{ijt} + \epsilon_{ijt} \quad (1)$$

where w_{ijt} is the logarithm of the hourly wage for individual i in group j at time t . k_j is the level of productive skills that is common to all members of group j , and is what we call “observable skill”. a_{ijt} is the level of productive skills for individual i of group j at time t ; we call this “unobservable skill”. r_t and ψ_t are the respective returns to k_j and a_{ijt} . ϵ_{ijt} represents a combination of measurement error, and other random determinants of wages.¹¹

For illustrative purposes, in this section we exclusively discuss identification of the return to college, but it will be clear that this general framework is appropriate for characterizing the identification of changes in discrimination.¹² In such a case, the two groups would be *BLACK* and *WHITE* workers (with otherwise identical observable characteristics). r_t would then be the discrimination coefficient while ψ_t would remain the return to unobservable skill.

The following assumptions about the terms in equation 1 are made:

1. $E[a_{ijt}] = a_j$
2. $Var[a_{ijt}] = \sigma_{aj}^2$
3. k_j and a_{ijt} is jointly independent of ϵ_{ijt}

¹¹One example of a model of wage determination that might motivate equation 1, is one where workers are paid their marginal productivity, where the aggregate production function is $Q_t(K_t, L_{1t}, \dots, L_{Nt}) = G_t(K_t) \cdot F\left(\sum_{i=1}^N L_{it} \cdot \exp\{r_t k_j + \psi_t a_{ij} + \epsilon_{ijt}\}\right)$.

¹²This framework is adopted in Chay (1996), who accounts for the potential biases caused by a changing return to unobserved skill, in his empirical evaluation of the impact of the 1964 Civil Rights Act on black-white relative wages. In that paper, however, identification of ψ_t is achieved in a different manner. It is done through shifts in conditional mean wages for white workers, as well as information contained in the residual autocovariance matrix of wages.

Assumptions 1 and 2 mean that we allow unobservable skill for an individual to vary over time, as long as its expected value and variance does not change over time. Note that we make no additional assumptions about the time series properties of a_{ijt} . Assumption 3 is meant to allow for the existence of another random determinant of earnings that may be completely unrelated to either observed or unobserved skill. For example, measurement error might be an example of such a random component. Special cases of the above set of assumptions include the case where residual earnings (from the mean for an observable group) is entirely due to skill ($\sigma_{\epsilon,t}^2 = 0 \forall t$), and where there is a permanent individual fixed effect ($a_{ijt} = a_{ij} \forall t$).

Assumptions 1 to 3 are not sufficient for identifying the relative return to college with the mean log wage of each education group. This is because in general $a_{COL} \neq a_{HS}$, so that

$$E[w_{i,COL,t}] - E[w_{i,HS,t}] = r_t + \psi_t (a_{COL} - a_{HS}). \quad (2)$$

(where we have normalized $k_{COL} - k_{HS} = 1$ to give r_t the interpretation of a relative return to college). The second term in equation 2 is the familiar bias encountered when estimating the return to college in the presence of nonrandom assignment of workers into education groups.

Furthermore, the magnitude of the bias can change over time, and it is directly proportional to $(\psi_t - 1)$. In particular,

$$\begin{aligned} & (E[w_{i,COL,t}] - E[w_{i,HS,t}]) - (E[w_{i,COL,1}] - E[w_{i,HS,1}]) \\ &= (r_t - r_1) + (\psi_t - 1)(a_{COL} - a_{HS}) \end{aligned} \quad (3)$$

where we make the normalization $\psi_1 = 1$, so that ψ_t is the return to unobserved productivity relative to the base year. This equation shows that the *change* in the return to college is

also unidentified, and the conventional estimate of the change — the change in the measured college-high school wage differential — is biased by the magnitude $(\psi_t - 1)(a_{COL} - a_{HS})$. In principle, any change in the college-high school differential could be due to a change in ψ_t , rather than r_t . It will thus be necessary to obtain information about ψ_t in order to identify $(r_t - r_1)$.

It is also necessary to obtain information about the magnitude of $(a_{COL} - a_{HS})$, the bias in period 1. Although there is a substantial empirical literature associated with attempts to account for this selection bias, there is little consensus on the size of this magnitude.¹³ In this study, our approach is to remain agnostic about the magnitude of the selection bias, but to focus on credible identification and estimation of ψ_t . Given ψ_t , we can then provide a range of estimates of the return to college that corresponds to the range of beliefs about the magnitude of $(a_{COL} - a_{HS})$.

2.3 Non-identifiability of the changes in the return to unobserved skill

The hypothesis that rising within-group wage dispersion represents an increase in the returns to unobserved skill raises the question of how one might estimate the magnitude of its increase over a certain period. In particular, how does the residual dispersion series in Figure 1 translate into changes in the return to unobservable skill? Using a single time series of statistics such as the residual standard deviation or the residual 90-10 percentile difference is problematic if there is any random component of wages which is unrelated to productive skills (e.g. measurement error, “luck”), and especially if this component’s variance changes over time.

¹³See Willis (1986) for an overview of the theoretical underpinnings of earnings functions and associated estimation problems. A summary of recent attempts to deal with the self-selection problem can be found in Card (1994).

More specifically, using equation 1, we have

$$Var[w_{ij1}] = \sigma_{wj,1}^2 = \sigma_{aj}^2 + \sigma_{\epsilon,1}^2 \quad (4)$$

$$Var[w_{ijt}] = \sigma_{wj,t}^2 = \psi_t^2 \sigma_{aj}^2 + \sigma_{\epsilon,t}^2$$

from which one can solve for ψ_t ,

$$\psi_t = \sqrt{\frac{\sigma_{wj,t}^2 - \sigma_{\epsilon,t}^2}{\sigma_{wj,1}^2 - \sigma_{\epsilon,1}^2}}. \quad (5)$$

It is clear that the within-group wage variances in two points in time are not sufficient to identify ψ_t . Even when the wage variance rises from period 1 to period t , it is possible for ψ_t to be less than 1, as long as $\sigma_{\epsilon,t}^2$ is sufficiently large relative to $\sigma_{\epsilon,1}^2$. That is, if the variance of the non-skill related component of wages rises, a rise in the observed wage variance implies less of a rise in the return to unobserved skill. By the same token, a fall in the wage variance from period 1 to t need not imply $\psi_t < 1$.

Note that by making the restriction $\sigma_{\epsilon,1}^2 = \sigma_{\epsilon,t}^2$, a rise in within-group wage variance unambiguously implies $\psi_t > 1$ (and obversely, a fall unambiguously implies $\psi_t < 1$). However, ψ_t can be made arbitrarily large (small) by setting $\sigma_{\epsilon,1}^2 = \sigma_{\epsilon,t}^2$ arbitrarily close to $\sigma_{wj,1}^2$ ($\sigma_{wj,t}^2$).

Thus we see that identification of the change in the return to unobserved skill depends crucially on the relative magnitudes of $\sigma_{\epsilon,1}^2$, $\sigma_{\epsilon,t}^2$ and σ_{aj}^2 . As an illustration of the potential range of ψ_t that results from various assumptions about these relative magnitudes, suppose we observe that $\sigma_{wj,1}^2 = .16$ and $\sigma_{wj,t}^2 = .2$, and that $\sigma_{\epsilon,1}^2 = \sigma_{\epsilon,t}^2$. With a normal distribution, this implies that the 90-10 wage differential is about 1 and 1.1 log points in period 1 and period t , respectively. Suppose first that the (residual) wage paid (i.e. the wage after subtracting out the conditional mean) is completely due to unobserved skill. Then it follows that $\psi_t = 1.12$,

the simple ratio of the standard deviations. On the other hand, suppose that the variance of unobserved skill were .01 in period 1, so that the 90-10 differential of unobserved skill (i.e. a_{ijt1} or a_{ijt}) were about .25 log points.¹⁴ Using equation 5, we would obtain $\psi_t = 2.23$.

2.4 Potential biases of conventional estimates

While it is clear that the various time series from Figure 1 are not sufficient to identify changes in the college premium, it is useful to consider what range of estimates would be obtained from various assumptions about 1) the magnitude of the selection bias in the base period and 2) the relative magnitudes of the variances of unobserved skill and other independent random factors.

Using the 1979 and 1991 data from Figure 1, Table 1 illustrates the range of estimates of changes in the college premium or discrimination between 1979 and 1991 that would arise under these various assumptions.¹⁵ The upper panel considers the case where the variance of ϵ_{ijt} is constant over time, while the lower panel considers a falling variance of ϵ_{ijt} over time (such that $\sigma_{\epsilon,t}^2 = .9 \cdot \sigma_{\epsilon,1}^2$). θ is defined as $\frac{\sigma_{\epsilon,1}^2}{\sigma_{w,1}^2}$, or the proportion of the 1979 wage variance which is attributable to “the non-skill” component of wages. λ is defined as $\frac{a_{COL} - a_{HS}}{E[w_{i,COL,1}] - E[w_{i,HS,1}]}$ (or $\frac{a_{BLACK} - a_{WHITE}}{E[w_{i,BLACK,1}] - E[w_{i,WHITE,1}]}$), the proportion of the initial period wage gap between the two groups that is attributable to differences in mean levels of unobservable skill. It parameterizes the “omitted-ability” bias in the initial period. The entries in the table are the estimates of the change in the college premium or discrimination coefficient which

¹⁴Note that we use ϵ_{ijt} to represent the combination of measurement error and other “random factors”; however, ϵ_{ijt} can also be thought of as another “skill” component that is independent of the joint distribution of k_j and a_{ijt} . For our purposes, what it is called is inconsequential to the present analysis, since we are only interested in the return to unobserved skill insofar as it affects estimates of changes in the return to observed skill.

¹⁵For the sake of exposition, we assume that the variance of wages are the same in both groups (*COL* and *HS*, or *BLACK* and *WHITE*). The possibility that they may vary across groups is instrumental in the identification strategy discussed later in the paper.

would arise under the corresponding assumption about λ , and θ (and the implied ψ_{91}). They are calculated from equations 3 and 5.¹⁶

The first thing to note from Table 1 is that whenever $\lambda = 0$ or $\psi_{91} = 1$ the conventional estimate of the college premium, the college-high school (regression-adjusted) wage gap, is an unbiased estimate of $(r_t - r_1)$. In other words, if either 1) the return to unobserved skill does not change over time or 2) the initial wage gap is an unbiased estimate of the true return to college, the change in this wage gap will be an unbiased estimate of the change in the return to college.¹⁷

The table illustrates that the 1979-1991 change in the college premium decreases, and the 1979-1991 change in discrimination increases, in both λ and θ . Furthermore, there is an interaction effect such that the greater is θ (and hence the greater is ψ_{91}), the more sensitive is the estimate of the change to the assumption about the magnitude of λ . For example, the third column of the upper panel of Table 1 shows that when $\theta = .5$ and $\psi_{91} = 1.17$, then estimates of the change in the college premium can range from .15 to .20, whereas if $\theta = .9$ and $\psi_{91} = 1.67$, the estimate can range between 0 and .2, depending on the assumption about the initial year's selection bias.

The lower panel points out that the degree of indeterminacy of the change in the college premium or discrimination increases if the variance of the non-skill component of wages falls over time. The final column of the lower panel shows that under the extreme assumption that all of the initial wage gap between blacks and whites are due to differences in unmeasured productivity, a doubling of the return to unobserved skill implies that the absolute magnitude

¹⁶More specifically, from equation 3, the entries are $(r_t - r_1) = \{(E[w_{i,COL,t}] - E[w_{i,HS,t}]) - (E[w_{i,COL,1}] - E[w_{i,HS,1}])\} \cdot [1 - (\psi_t - 1)\lambda]$.

¹⁷ $(a_{COL} - a_{HS})$ also needs to remain constant over time. This issue will be further discussed in section 4.

of the discrimination coefficient *falls* by about .09 log points rather than increasing by .04 log points, which is suggested by the data when the returns to unobserved skill are ignored.

The table suggests that identifying ψ_t will be crucial to placing lower and upper bounds on the change in the return to college or the change in discrimination. The next section discusses an identification strategy that utilizes the variability in within-group wage variances.

3 Identification of changes in the return to unobserved skill

We show in this section that across-group variation in within-group wage variances allows identification of ψ_t . The identification strategy has the advantage that it does not require the full specification of the time series properties of the errors as defined in equation 1. It also accommodates a non-stationary “noise” component. We describe identification under two different assumptions about the “noise” or “non-skill” component of earnings. We then discuss over-identification of ψ_t when more than two years of data are available.

3.1 Case 1: constant “noise” variance over time

Consider the case where, in addition to assumptions 1–3, we make the restriction that $\sigma_{\epsilon,t}^2 = \sigma_{\epsilon,s}^2 \forall t, s$.¹⁸ From equation 4, we obtain

$$\sigma_{wj,t}^2 = \sigma_{\epsilon,1}^2 (1 - \psi_t^2) + \psi_t^2 \sigma_{wj,1}^2. \quad (6)$$

In principle, if population within-group wage variances were known, ψ_t is identified as long as there are at least two groups with differing wage variances in the initial period. $\sigma_{\epsilon,1}^2$ is also identified in this case. Period t ’s wage variance is a linear function of the period 1

¹⁸In conjunction with assumption 3, this restriction implies $Var[\epsilon_{jt}] = Var[\epsilon_{ks}] \forall j, k, t, s$. Note that assumption 3 is sufficient but not necessary. More generally, we only require that $Var[\epsilon_{ijt}|j] = \sigma_{\epsilon,t}^2$, $Cov[\epsilon_{ijt}, a_{ijt}] = 0 \forall j$, and $E[\epsilon_{ijt}|j] = \mu_\epsilon$.

variance. Note that we make no assumptions about $Cov[\epsilon_{ij1}, \epsilon_{ijt}]$ or $Cov[a_{ij1}, a_{ijt}]$. The joint distributions of $(\epsilon_{ij1}, \epsilon_{ijt})$ and (a_{ij1}, a_{ijt}) need only satisfy assumptions 1–3 and the constancy of the variance of ϵ_{ijt} over time.

3.2 Case 2: non-constant “noise” variance over time

Here we relax the restriction $\sigma_{\epsilon,t}^2 = \sigma_{\epsilon,s}^2 \forall t, s$. From equation 4, we obtain

$$\sigma_{wj,t}^2 = \left(\sigma_{\epsilon,t}^2 - \psi_t^2 \sigma_{\epsilon,1}^2 \right) + \psi_t^2 \sigma_{wj,1}^2. \quad (7)$$

Again, two different within-group wage variances in the initial period are sufficient to identify ψ_t . Even though $\sigma_{\epsilon,1}^2$ and $\sigma_{\epsilon,t}^2$ are now both unidentified, the sign of the intercept in the above linear function, tells us about the relative proportional growth of the scale of ϵ_{ijt} versus the scale of a_{ijt} .

3.3 Over-identification of ψ_t using multiple years

In principle, ψ_t is over-identified in equations 6 and 7 whenever there are more than two different wage variances in either of the periods. Given a sufficiently large number of wage observations per group, we can appeal to large sample theory and use the asymptotic normality of the estimated variance for each group in order to construct an appropriate test statistic of the hypothesized linear restriction of equation 6 or 7.

In practice, however, we may have a relatively small number of observations per group, but a relatively large number of groups. In this case, we can treat each group as an observation, and consider the limiting sampling distribution of estimators of ψ_t as the number of groups becomes large.

Equation 7 is equivalent to

$$\hat{\sigma}_{wj,t}^2 = \left(\sigma_{\epsilon,t}^2 - \psi_t^2 \sigma_{\epsilon,1}^2 \right) + \psi_t^2 \hat{\sigma}_{wj,1}^2 - \psi_t^2 u_{j1} + v_{jt}. \quad (8)$$

where u_{jt} and v_{jt} are defined as the mutually uncorrelated sampling errors of $\sigma_{wj,1}^2$ and $\sigma_{wj,t}^2$, respectively; hats denote sample estimates. Randomness in this model comes from two sources. First, σ_{aj}^2 is drawn from a distribution of skill variances. Second, there is the randomness that arises from estimation of the within-group variances. This generates an independent but not identically distributed sequence of pairs $(\hat{\sigma}_{wj,1}^2, \hat{\sigma}_{wj,t}^2)$. Equation 8 represents a typical bivariate errors-in-variables linear regression framework, and OLS will thus give inconsistent estimates of ψ_t^2 . We know that the OLS estimate is

$$\widetilde{\psi}_t^2 = \frac{\widehat{Cov} [\hat{\sigma}_{wj,1}^2, \hat{\sigma}_{wj,t}^2]}{\widehat{Var} [\hat{\sigma}_{wj,1}^2]} \xrightarrow{p} \psi_t^2 \cdot \left(\frac{Var [\sigma_{wj,1}^2]}{Var [\sigma_{wj,1}^2] + \frac{1}{J} \sum_{j=1}^J Var [u_{j1}]} \right) \quad (9)$$

where the hats denote estimates from a sample in which each observation is an estimated variance from one of J groups. Probability limits here are defined as J , the number of groups, goes to infinity.

Consistent estimation of $Var [\sigma_{wj,1}^2]$ will allow at least one consistent estimator of ψ_t^2 .

For example, we can use

$$\widetilde{Cov} [{}_a \hat{\sigma}_{wj,1}^2, {}_b \hat{\sigma}_{wj,1}^2] \xrightarrow{p} Var [\sigma_{wj,1}^2] \quad (10)$$

where the subscripts a and b denote two independent samples (which we can make by randomly splitting the sample of all wage observations in two halves). We can use this result to form the consistent estimator

$$\widehat{\psi}_t^2 = \frac{\widehat{Cov} [\hat{\sigma}_{wj,1}^2, \hat{\sigma}_{wj,t}^2]}{\widetilde{Cov} [{}_a \hat{\sigma}_{wj,1}^2, {}_b \hat{\sigma}_{wj,1}^2]} \xrightarrow{p} \psi_t^2. \quad (11)$$

An alternative estimator for ψ_t^2 can be formed using data from a third period s . It is easy to show that

$$\widehat{\psi_t^2} = \frac{\widehat{Cov}[\widehat{\sigma}_{wj,t}^2, \widehat{\sigma}_{wj,s}^2]}{\widehat{Cov}[\widehat{\sigma}_{wj,1}^2, \widehat{\sigma}_{wj,s}^2]} \xrightarrow{p} \psi_t^2 \quad (12)$$

provided that the sampling error of the estimated within-group variances in any period is uncorrelated with the sampling error in estimates from any other period (e.g. u_{j1} and v_{jt}) and the true within-group variance in any period is uncorrelated with the sampling error from any year.

Finally, it is straightforward to show that

$$\widehat{\psi_{s,1}^2} \widehat{\psi_{t,s}^2} \xrightarrow{p} \psi_t^2 \quad (13)$$

where $\widehat{\psi_{s,1}^2}$ is a consistent estimate for $\psi_{s,1}^2$, which is the slope coefficient when $\sigma_{wj,s}^2$ is expressed as a linear function of $\sigma_{wj,1}^2$ (and similarly for $\widehat{\psi_{t,s}^2}$). Therefore, with multiple years of data, ψ_t is over-identified and an appropriate test statistic of the hypothesis of equality of the alternative estimates of ψ_t can be constructed.

Note that none of the identification strategies discussed above rely on a full specification of the time-series properties of the error components. In particular, serial correlation over time of either a_{ijt} or ϵ_{ijt} is unrestricted.¹⁹ We only require that the 3 assumptions of section 2 hold.

4 Between- vs. Within-cohort data from the CPS

In this section, we briefly discuss two alternative ways in which CPS micro-data on earnings can be used to implement our identification strategy. In particular, we contrast using

¹⁹An alternative identification strategy would involve specifying a parametric form of the time series process of a_{ijt} and ϵ_{ijt} , and using the autocovariances of the wage residuals to recover ψ_t . For example, see Card and Lemieux (1994) and Chay (1996).

longitudinal wage (within-cohort) data with using wage data from a group of workers with similar observable characteristics in two different time periods (between-cohort). We discuss the advantages and disadvantages of both approaches, attempting to highlight the pitfalls of each that should be kept in mind when interpreting our results. We also describe the relevant features of CPS earnings data which complicate our empirical analysis.

4.1 Between- vs. Within-cohort comparisons

Even when the return to unobserved skill is ignored, when researchers infer from a rising college-high school wage gap that the return to college is rising, they implicitly make an assumption similar in spirit to assumption 1 of section 2.²⁰ Roughly speaking, this is the assumption that the unobserved “quality” difference between college and high school-educated workers remains the same in the two years. If this assumption were not made, then equation 3 would more generally be

$$\begin{aligned}
& (E[w_{i,COL,t}] - E[w_{i,HS,t}]) - (E[w_{i,COL,1}] - E[w_{i,HS,1}]) \\
&= (r_t - r_1) + (\psi_t - 1)(a_{COL,1} - a_{HS,1}) \\
&\quad + \psi_t \{(a_{COL,t} - a_{HS,t}) - (a_{COL,1} - a_{HS,1})\}
\end{aligned} \tag{14}$$

where $a_{COL,1}$, for example, is the mean level of unobservable skill for college-educated workers in period 1. Therefore, even the calibration exercise in Table 1 relies on the assumption that $(a_{COL,t} - a_{HS,t}) = (a_{COL,1} - a_{HS,1})$.

This restriction is often thought to be justified when longitudinal data is used. That

²⁰Strictly speaking, assumption 1 is sufficient but not necessary in our framework. We only need $E[a_{ijt}] - E[a_{ikt}] = \Delta a_{jk} \ \forall t$ (the mean unobserved skill gap to remain constant over time).

is, with repeated wage observations of the same individuals over time, one might think that the mean college-high school unobserved skill gap ($a_{COL} - a_{HS}$) is constant for a fixed group of individuals over time. In fact, true longitudinal data such as the NLSY or PSID is unnecessary if information from the autocovariances of residual wages remains unused, as it is in our identification strategy. As Deaton (1985) points out, by following fixed birth-year cohorts with repeated cross-sections of micro-data, one can construct synthetic panel data, which become a viable alternative to real longitudinal data, which often have small sample sizes and are unrepresentative of the population of interest. In fact, the independence of the samples over time is essential for estimation in our application of the classical errors-in-variables framework described in section 3.

The identifying assumption that ($a_{COL} - a_{HS}$) is constant over time is untestable, whether in a within-cohort or between-cohort context. However, it is still useful to examine demographic characteristics of the population of interest during the period 1979-1991 when forming a judgment about the validity of this assumption. In particular, we argue that if the distribution of educational attainment changes significantly over time, the constancy of ($a_{COL} - a_{HS}$) over time is less likely to hold.

Table 2 summarizes the educational distributions of white and black male earners in 1979 and in 1991. By race, and by five-year experience cohorts, we report the distribution of education into four educational groups (<12, 12, 13-15, 16 or more) in 1979 and in 1991. The third columns for both white and black earners are the educational distributions when following a fixed cohort of workers (i.e. the educational distributions for workers that have 12 more years of experience in 1991 than they did in 1979).²¹

²¹Of course, there may be some within-cohort changes here due to 1) individuals obtaining more schooling

As expected, within-cohort educational distributions look very similar in the four oldest cohorts, for both white and male workers. On the other hand, the between-cohort comparisons show that there are significant differences in educational distributions. In particular, for the oldest two cohorts of white males, and for all of the black male cohorts, the 1991 between-cohort comparison groups have higher levels of education. For example, about 10 percent of black earners with 11-15 years of experience in 1991 were high school dropouts and 20 percent were college graduates, while among those with 11-15 years of experience in 1979, 23 percent were dropouts and 13 percent were college graduates. From these summary statistics, we argue that the assumption of an unchanging unobserved skill gap, $(a_{COL} - a_{HS})$ or $(a_{BLACK} - a_{WHITE})$, may be more credible for within-cohort than for between-cohort comparisons.²²

One shortcoming of a within-cohort analysis, however, is that it may be complicated by the possibility of an interaction between education and the return to experience. In a within-cohort context, equation 3 is more precisely written as

$$\begin{aligned} & (E[w_{i,COL;OLD,t}] - E[w_{i,HS;OLD,t}]) - (E[w_{i,COL;YNG,1}] - E[w_{i,HS;YNG,1}]) \\ &= (r_t^{OLD} - r_1^{YNG}) + (\psi_t - 1)(a_{COL} - a_{HS}) \end{aligned} \quad (15)$$

where *OLD* and *YNG* denote the age of the individuals in the present period. While $(r_t^{OLD} - r_1^{YNG})$ may sometimes be the parameter of interest, we are often also interested in $(r_t^{YNG} - r_1^{YNG})$ or $(r_t^{OLD} - r_1^{OLD})$. Only when the return to a year of experience is equal among both educational groups, at each point in time, does the within-cohort analysis yield

(especially for the youngest cohort, 1-5 years) and 2) nonrandom selection out of the workforce, since the sample is the population of workers with nonzero wages.

²²We also argue that 2 may be more plausible in a within-cohort than in a between-cohort context, for the same reasons.

$(r_t - r_1)$.

In this study, our approach is to report results from both between- and within-cohort analyses, while cautioning that the interpretation of our results should be tempered by two considerations. One is that the identifying assumption of the constancy of $(a_{COL} - a_{HS})$ may be more credible in a within-cohort context, and the other is that a between-cohort analysis may be more appropriate for identifying $(r_t - r_1)$ in the presence of significant education-experience interaction effects.²³

Since our analysis essentially “adjusts” conventional (reduced-form) estimates to take into account the return to unobserved skill, we first provide a brief summary of the between- and within-cohort changes in college-high school and black-white wage differentials during the 1980s, as calculated from CPS earnings data.

Figure 2 graphs the change in the college-high school wage differential for four separate experience cohorts.²⁴ The first thing to note from the figure is that the change in the college-high school differential during the 1980s varies by experience cohort.²⁵ For example, the 1979-1991 change is as low as .12 for the between-cohort series for cohort 4, and as high as .24 for cohort 1. Second, within-cohort changes during the 1980s seem to be similar in direction and magnitude to between-cohort changes, with the possible exception of cohort 1. Panel A shows that by 1991, among those who had 6-10 years of experience in 1979, the college-high school wage differential widened somewhat less (by about .07 or .08 log points)

²³When examining changes in discrimination during the 1980s, we may be equally interested in $(r_t^{OLD} - r_1^{YNG})$, as this has the interpretation of the change in discrimination for a particular birth-year cohort during the decade.

²⁴See the note to Table 2 for details on the data. Charts are constructed from the coefficients from the following regression: $\log(\text{wage})$ on indicator variables for each year of experience, education category (<12,13-15,16+) and race.

²⁵The youngest cohort we examine are those with 6-10 years of potential experience. An examination of within-cohort changes in educational distributions suggested that among those with 5 or less years of experience, many were still in the process of obtaining more education.

than the differential for those with 6-10 years of experience in 1991.

Figure 3 depicts the well-documented stagnation of black-white wage convergence during the 1980s. In particular, we see that both between- and within-cohort, blacks in cohort 1 lost about .05 log points relative to their white counterparts during the 1979-1991 period. For cohorts 2-4, there was little movement in the black-white wage differential during the decade. Figure 4 decomposes Panel A into two education groups, those with high school or less and those with at least some college attendance. It shows that the decline in the black-white differential was more significant for young, relatively more-educated workers. This happened both between- and within-cohort.

Finally, Figure 5 summarizes movements in residual, or within-group wage dispersion during the 1980s.²⁶ It shows that on the whole, within-group wage inequality increased by about the same amount across all four cohorts, by about .04 or .05 log points; within-cohort residual dispersion seems to have grown slightly more than between-cohort dispersion for all cohorts throughout the decade.²⁷

4.2 Top-coding in the CPS

While CPS earnings data constitute a large and representative sample of workers in the United States, and hence seems especially ideal for our identification strategy, its major drawback is that its earnings data are censored. The CPS public use data “top-codes” weekly earnings at \$999 in 1979 until 1989 and later, where the top-code rate is at \$1999.

Censoring is especially problematic in the estimation of equation 8, since identification comes

²⁶The residual standard deviation is the root mean squared error of a regression of $\log(\text{wage})$ on a fully-interacted set of dummy variables for experience, education (4 categories) and race.

²⁷Note that, as in Figures 2-4, no correction is made for the top-coding procedures used by the CPS public use data. Since the top-code rate changed from \$999 to \$1999 between 1988 and 1989, one should expect to see a slight jump in the wage differentials and in the residual standard deviation between 1987 and 1989, even in the absence of any real change.

from across-group variation in within-group variances.

Figure 6 summarizes the extent of earnings censoring in our data. As expected, it shows the fraction of men who are at the censoring point increasing steadily, as nominal wages rise over time, between 1979 and 1988; it drops off sharply at 1989. As Panel B shows, censoring can be quite serious in the couple of years immediately preceding 1989; the top-code rate is about .3 for the oldest cohorts. Fortunately, the censoring problem is not serious for at least two of the years, 1979 and 1989. This is especially fortuitous, since we are mainly interested in the long-term changes in the return to unobserved skill during the 1980s. On the other hand, it is clear that if data from the mid-80s are used as a source of over-identification, as proposed in section 3.3, an alternative estimator of the within-group wage variance will be necessary, since the usual sample estimate of the variance will likely be significantly downward biased.

5 Estimation and Results

In this section, we report the results of several alternative procedures used to estimate ψ_t , the return to unobserved skill (relative to the base year). Each of our approaches utilizes equation 7 — that is, across-group variation in within-group wage dispersion over time — in order to estimate ψ_t .

5.1 Estimation using aggregated data

Our first approach makes use of the fact that ψ_t is identified when only two groups' wage variances are known in two separate periods. In order to make use of the large overall sample size of our data, and hence reduce the sampling error of the estimated moments needed to

estimate ψ_t , we divide the sample into two arbitrary groups, and then calculate a “weighted-average” within-group wage variance for each of these groups for each period.

More formally, we have a total of J different groups, where each group is uniquely identified by the combination of 1) years of schooling, 2) labor market experience, and 3) race. Divide the J groups into two sets, containing J_1 and J_2 groups, respectively. From equation 7 we can obtain:

$$\sigma_{wJ_1,t}^2 = \left(\sigma_{\epsilon,t}^2 - \psi_t^2 \sigma_{\epsilon,1}^2 \right) + \psi_t^2 \sigma_{wJ_1,1}^2 \quad (16)$$

and

$$\sigma_{wJ_2,t}^2 = \left(\sigma_{\epsilon,t}^2 - \psi_t^2 \sigma_{\epsilon,1}^2 \right) + \psi_t^2 \sigma_{wJ_2,1}^2. \quad (17)$$

where $\sigma_{wJ_1,t}^2$ is

$$\frac{1}{\sum_{j=1}^{J_1} N_j} \sum_{j=1}^{J_1} N_j \sigma_{wj,t}^2 \quad (18)$$

and similarly for $\sigma_{wJ_2,t}^2$. N_j is the number of observations in group j . Solving for ψ_t^2 , we obtain

$$\psi_t^2 = \frac{\sigma_{wJ_2,t}^2 - \sigma_{wJ_1,t}^2}{\sigma_{wJ_2,1}^2 - \sigma_{wJ_1,1}^2} \quad (19)$$

This equation illustrates that as the *difference* in two groups’ (average) within-group wage variance expands over time, the implied return to unobserved skill rises.²⁸

The choice of how to aggregate the groups is arbitrary, but a natural starting point is to define the two groups by education category. Table 3 presents the “weighted-average” within-group wage variances for workers with 12, and 16 or more years of schooling, and between 6 and 25 years of experience in 1979 and 1989 (and for the within-cohort comparison, those

²⁸A finding that one group’s wage variance is smaller than the other in one period, but larger in the next period is completely inconsistent with our model. (A negative value of ψ_t^2 would be obtained).

with 16-35 years of experience in 1989).²⁹ We choose these two years because the top-coding in these years is negligible. The table shows that within-group wage variances vary by education category. In 1979, the average residual variance for those with 16 or more years of schooling is about .05 log points (or .22 log points if comparing standard deviations) greater than those with 12 years of schooling; this seems to generally hold when the sample is divided into four 5-year cohorts as well.

As equation 19 suggests, an increase in the return to unobserved skill implies that the college-high school *variance* differential should expand. Table 3 shows that overall, the differential does not rise significantly between-cohort; in fact, it contracts slightly for the oldest two cohorts. On the other hand, the within-cohort column shows that the gap expands in three of the four cohorts. The third columns for the between- and within-cohort comparisons report the implied ψ_{89} given the appropriate four estimates of within-group wage variances.³⁰

For workers with 6 to 25 years of experience, the between-cohort estimate is a 2 percent rise, while the within-cohort estimate is about an 11 percent increase in the return to unobserved skill, although both have standard errors of about .06. The implied estimates by the four 5-year cohorts show some differences, with two of the cohorts exhibiting increasing returns to skill and the other showing decreasing returns during the 1979-1989 period. As expected, however, the estimated sampling variability increases significantly when looking at the smaller cohorts.

The bottom panel of Table 3 replicates the analysis for the case where the two groups are not defined by education, but instead by race. It shows that the black-white gap in within-

²⁹The sample analogue of 18 is simply the residual variance in a regression of log wages on fully interacted sets of single-year experience, single-year education, and race indicator variables.

³⁰The asymptotic standard errors are calculated using a first-order Taylor series approximation of equation 19.

group wage variances is not as pronounced as the college-high school gap. The average wage variance for white workers are slightly higher than those for blacks in 1979 (the exception being for those with 16-20 years of experience in 1979). Furthermore, the gaps expand both between-cohort and within-cohort during the 1980s. However, since the initial year variance differential is small and the sample of blacks is relatively small, these moments are less useful in identifying ψ_{89} , as the estimated sampling errors are relatively large.³¹

5.2 Estimation using disaggregated data

One drawback of the above estimation procedure is that the aggregation of the J cells into two groups is arbitrary. As Table 3 shows, a division along the lines of education appears to be more useful than a grouping by race, but it is not obvious which aggregation scheme will work best to precisely estimate ψ_t . An alternative approach that uses the information on within-group wage variances is to treat each of the J groups as a single observation, as discussed in section 3.3.

Tables 4A and 4B provide summary statistics of the “cell-level” data used for the subsequent between- and within-cohort analysis. We exclude cells that contain fewer than 30 observations, those cells in which there is at least one top-coded observation below the cell median, and those cells in which more than 5 percent are earning below the federal minimum wage (plus 5 cents/hour). For the within-cohort cell-level data, we exclude those with less than 6 years of potential experience in 1979.

When making our sample restriction, we attempt to retain as many cells as possible, while eliminating cells for which estimation of the true wage variance requires too many

³¹For the 16-20 cohort, an estimate for ψ_{89} could not be calculated, since the estimated ψ_{89}^2 was negative.

assumptions on the distribution of wages. In particular, the erosion of the real value of the minimum wage during the 1980s may cause within-group wage variances to rise among low-wage workers, independently of any change in the return to unobserved skill. In principle, we could explicitly model the effect of the minimum wage on the distribution of wages within our model of observable and unobservable skill, and subsequently attempt to estimate what wage variances would be in the absence of the minimum wage. However, that is beyond the scope of this paper, and we thus focus on workers who are largely unaffected by the minimum wage.

Second, since the substantial top-coding that occurs in the 1980s causes sample variances to underestimate true within-group wage variances, we use an alternative estimator of a cell variance that is based on using the “lower tail” of the empirical distribution of wages. This estimator is simply the average of the squared deviations from the median for all observations below the median.³² We consider cells for which some top-coded observations are below the median to be inappropriate for both the usual sample variance and this “lower tail” estimator.

Table 4A shows that there are a total of 503 cells that appear at least once in the 7 years of data that are used.³³ The number of valid cells falls steadily until 1987, due to highly censored cells being dropped. Table 4B reports the summary statistics for the within-cohort data. Over time, workers age and hence leave the sample (wince we keep only those between 18 and 64 years of age). Hence the number of cells declines over time.

³²If we assume symmetry of the within-group wage distribution, this will be a consistent estimate of the within-group variance.

³³We use every other year, since the CPS’s overlapping rotation scheme creates nonindependent samples in consecutive years. The independence of the sampling errors over time is crucial when using adjacent years’ wage variances as instruments.

How did within-group wage variances covary over the course of the 1980s? Figure 7A depicts the empirical relation between within-group wage variances in 1989 and the variances in 1979 for the between-cohort data. It plots the data to be used for estimation of the simple linear relation represented in equation 7. As the figure shows, there appears to be a positive correlation between within-group variances in 1979 and the within-group wage variances of workers with identical observable characteristics in 1989. It appears as though there is sufficient variation in within-group wage variances to identify the slope parameter in the hypothesized relation described in equation 7.

The plotting symbols provide more information on how the slope parameter is identified. The circles, squares, and triangles denote those cells with individuals with 12, 13-15, and 16 or more years of schooling, respectively. The plot shows that the magnitude of the within-group variance generally rises with more education. Within our framework, this implies that there is a greater variance in unobserved skill among those with more years of schooling. The observable implication of an increase in the return to unobservable skill ($\psi_{89} > 1$) is a cell's wage variance should grow faster between 1979 and 1989, the larger is its variance in 1979.

The dashed line represents our estimate of the intercept and slope of the linear relation between the two years' variances.³⁴ The estimate of the slope is 1.233, which implies that $\psi_{89} = 1.110$, an 11 percent increase in the return to unobserved skill during the decade. The estimated intercept is -.00288. If the non-skill component of wages is assumed to have a constant variance, as in equation 6, this implies that $\sigma_{\epsilon}^2 = .0124$; the mean within-group variance in 1979 is .168, which means that about 7.3 percent of the within-group wage variance is attributable to the non-skill error component. More generally, when the non-skill

³⁴In the next section, we describe in detail how we obtain consistent estimates of these parameters.

error component is non-stationary, as in equation 7, the intercept gives us no information about the relative magnitudes of the skill and the non-skill error variances.

Figure 7B is the analogous plot for the within-cohort comparison. The plot looks very similar to Figure 7A, exhibiting a significant positive correlation between the two years' variances. The slope of the relation is slightly steeper, at 1.425, implying $\psi_{89} = 1.194$, about a 20 percent increase in the payoff to unobservable skill. Again, if we assume equation 6 is the underlying relation, the estimated intercept of -.0169 implies $\sigma_\epsilon^2 = .0397$; at the average within-group variance of .163, this would imply that about 24 percent of the within-group wage variance in the initial period is due to the independent non-skill component.

5.2.1 Instrumental variable estimation

As noted in section 3.3, since each cell's estimated variance is subject to sampling error, OLS will yield an inconsistent estimate of the slope of the hypothesized linear relation of the two years' variances. In order to consistently estimate ψ_{89} , we randomly split the entire sample into two equally sized subsamples, and use the one subsample's estimates of the cell variances as an instrument for the other subsample's estimates. More formally, from equation 8, we have

$$\begin{aligned}\hat{\sigma}_{wj,89}^2 &= \left(\sigma_{\epsilon,89}^2 - \psi_{89}^2 \sigma_{\epsilon,79}^2\right) + \psi_{89}^2 \hat{\sigma}_{wjs,79}^2 - \psi_{89}^2 u_{js79} + v_{js89} \\ \hat{\sigma}_{wj,79}^2 &= \left(\sigma_{\epsilon,79}^2 - \frac{1}{\psi_{89}^2} \sigma_{\epsilon,89}^2\right) + \frac{1}{\psi_{89}^2} \hat{\sigma}_{wjs,89}^2 - \frac{1}{\psi_{89}^2} v_{js89} + u_{js79}\end{aligned}\tag{20}$$

where the subscript s identifies which subsample for each cell is used; terms without an s subscript are for estimates (and errors) in which the sample is not split into the two subsamples. The second equation is just the reverse regression where the dependent variable is instead the within-group wage variance in 1979. In order to use all the available information

from the two years' data, we simultaneously estimate both the slope (ψ_{89}^2) from the first equation and its reciprocal ($\frac{1}{\psi_{89}^2}$) from the reverse regression.

If we denote s' as the complementary subsample of s , then it is clear that $\hat{\sigma}_{wjs',79}^2$ and $\hat{\sigma}_{wjs',89}^2$ are valid instrument for $\hat{\sigma}_{wjs,79}^2$ and $\hat{\sigma}_{wjs,89}^2$, respectively. In order to simultaneously estimate the two linear relations of equation 20, we simply define $\left[1(1st) \cdot \hat{\sigma}_{wj,89}^2 + 1(2nd) \cdot \hat{\sigma}_{wj,79}^2\right]$ as the dependent variable, $\left[1(1st), 1(1st) \cdot \hat{\sigma}_{wjs,79}^2, 1(2nd), 1(2nd) \cdot \hat{\sigma}_{wjs,89}^2\right]'$ as the 4×1 vector of regressors, and $\left[1(1st), 1(1st) \cdot \hat{\sigma}_{wjs',79}^2, 1(2nd), 1(2nd) \cdot \hat{\sigma}_{wjs',89}^2\right]'$ as the 4×1 vector of instruments, where $1(\cdot)$ is an indicator variable for the equation to which the observation belongs. For the between-cohort estimate, we use 1324 observations (331 distinct cells $\times$ 2 subsamples per group $\times$ 2 linear relations). In estimation, we weight each cell by its combined sample size from both years.

The first row of Table 5A reports this procedure's estimates of the slopes in both of the linear relations in equation 20. The slope of the 1989 variance as a function of the 1979 variance is estimated to be 1.12. The inverse relation gives an additional estimate of $\frac{1}{.746} = 1.34$. The errors are, in general, heteroskedastic (even after weighting by the square root of the sample size), and are by construction correlated across observations. Thus we report robust standard errors which allow for unrestricted heteroskedasticity, and intra-cluster correlation of the errors, where each of the J groups is a separate cluster. Given these two reduced-form estimates, we can test the over-identifying restriction implied by our model, while simultaneously providing a more efficient estimator of ψ_{89} under the assumption that the restriction is true.

We do this by minimizing the quadratic form

$$\left(1.12 - \widehat{\psi}_{89}^2, .746 - \frac{1}{\widehat{\psi}_{89}^2}\right) \widehat{V}^{-1} \left(1.12 - \widehat{\psi}_{89}^2, .746 - \frac{1}{\widehat{\psi}_{89}^2}\right)' \quad (21)$$

with respect to $\widehat{\psi}_{89}^2$, where $\widehat{V}$ is the estimated robust covariance matrix of the reduced-form estimates. As Table 5A shows, this minimum distance estimate is 1.11, or about an 11 percent rise in the return to unobserved skill between 1979 and 1989. The resulting goodness of fit statistic, which is asymptotically distributed as $\chi^2(1)$, is 1.88. The first row of Table 5B reports the analogous results for the within-cohort data. Here ψ_{89} is estimated to be 1.194 with an estimated standard error of .058; the goodness of fit statistic is .009.

A simple, alternative to the “split-sample” estimator described above, is an instrumental variable estimator which uses within-group variances from another year as an instrument for the within-group variance in 1979. This approach utilizes the information from the other years, and takes advantage of the independence of the CPS samples over time. The first two columns of Table 5A report the resulting reduced form slope estimates as well as the implied estimate of ψ_{89} , when each of the years’ data is used as the instrument.³⁵ They show that these alternative estimates are similar in magnitude to the split-sample estimate. Each of the estimates that use the 1985 and 1987 data as instruments is slightly larger than the other estimates, but their sampling errors are also significantly larger. The final row reports two-stage least squares estimates where the 1981 and 1991 variances are both used as instruments. The resulting estimate of ψ_{89} is similar in magnitude to the other alternative estimates.

³⁵Note that even though many of the cell variance estimates will be underestimated due to top-coding during the mid-1980s, as long as we maintain the assumption that each cell estimate’s expected value (and its sampling error) is jointly independent of the sampling errors in equation 20, the conventional sample variance estimates still comprise a valid instrument.

The final two columns of Table 5A report the analogous estimates that result if the “lower tail” estimates of the variances are the instruments. Even though Table 4A and Table 4B suggest that this variance estimate is overestimating the true variance (based on the data from 1979 and 1989), they are still valid instruments for the same reasons that the conventional variance estimates are valid. The resulting estimates of ψ_{89} look very similar to the estimates of the second column.

The rest of Table 5B reports the analogous instrumental variable estimates for the within-cohort data. Overall, the various alternative estimates are slightly larger than shown by the between-cohort data, with many of estimates within a standard error of 1.20. Note that the number of available cells for estimation is significantly smaller than for the corresponding between-cohort estimate. The estimated standard errors suggest that the within-cohort estimates of ψ_{89} are less precisely estimated.

5.2.2 Minimum distance estimation

Finally, we exploit all of the over-identifying restrictions discussed in section 3.3 to obtain a measure of how well our parsimonious model of unobservable skill fits the data on within-cell variances in the 1979-1991 period. We do this in a method-of-moments framework, whereby we fit the elements of the theoretical autocovariance matrix of the estimated within-group variances, as implied by our underlying structural model, to the empirical autocovariance matrix.

Given T periods, we let M be the $T \times T$ autocovariance matrix of the estimated within-group wage variances. After adding sampling errors to equation 4, we obtain that the lower

triangular of M is

$$\begin{pmatrix} Var[\sigma_{aj}^2] + \sigma_{u1}^2 & & & \\ \vdots & \ddots & & \\ \psi_{T-1}^2 Var[\sigma_{aj}^2] & \cdots & \psi_{T-1}^4 Var[\sigma_{aj}^2] + \sigma_{uT-1}^2 & \\ \psi_T^2 Var[\sigma_{aj}^2] & \cdots & \psi_{T-1}^2 \psi_T^2 Var[\sigma_{aj}^2] & \psi_T^4 Var[\sigma_{aj}^2] + \sigma_{uT}^2 \end{pmatrix} \quad (22)$$

where for notational purposes, we have that $\sigma_{ut}^2 = \frac{1}{J} \sum_{j=1}^J Var[u_{jt}]$, the average variance of the sampling error across all cells for year t . We are thus using $\frac{T(T-1)}{2}$ moments to identify $2T$ parameters; the model is hence over-identified whenever 4 or more years of data are used.

We minimize the criterion function

$$\left(Vec[\widehat{M}] - Vec[M] \right)' \left(Vec[\widehat{M}] - Vec[M] \right) \quad (23)$$

with respect to the parameters in 22 where $\widehat{M}$ is the empirical autocovariance matrix of the cell variances.³⁶ Altonji and Segal (1994) present evidence from Monte Carlo experiments that suggest that optimal minimum distance estimation of covariance structures can be potentially seriously biased in small samples. Thus, we choose the identity matrix as the weighting matrix in the criterion function.

Since we use data from years in which there is a nontrivial degree of top-coding, we use the “lower tail” estimates of the cell variances in this procedure. The summary statistics for the variances in 1979 and 1989 from Tables 4A and 4B suggest that the assumption of a strictly symmetrical within-group wage distribution, which justifies using the lower-tail estimate, does not hold in these data. However, we can still treat the lower-tail estimate as a proxy for the variance.

³⁶The variances and covariances in $\widehat{M}$ are weighted variances and covariances, where the weights are the sum of the number of observations in 1979 and 1989.

We justify using this estimate as a valid proxy by making the assumption that

$$\sigma_{wj,t}^{2'} = \alpha + \beta \sigma_{wj,t}^2 + \omega_{j,t} \quad (24)$$

where $\sigma_{wj,t}^{2'}$ is the population analogue of the lower-tail estimate, and $\omega_{j,t}$ is a specification error that is uncorrelated with the true variance $\sigma_{wj,t}^2$ and is uncorrelated with $\omega_{j,s} \forall s \neq t$. As a result, $Var[\hat{\sigma}_{wj,t}^{2'}] = \beta^2 Var[\sigma_{wj,t}^2] + Var[\omega_{j,t}]$, and $Cov[\hat{\sigma}_{wj,t}^{2'}, \hat{\sigma}_{wj,s}^{2'}] = \beta^2 Cov[\sigma_{wj,t}^2, \sigma_{wj,s}^2] \forall s \neq t$.³⁷ The theoretical autocovariance matrix of the lower-tail estimate, will be $\beta^2 M$, except that the set of nuisance parameters $\sigma_{u,t}^2$ will have a different interpretation.³⁸ Figures 8A and 8B plot the lower-tail estimate against the conventional variance estimate for 1979 and 1989, the two years in which the degree of top-coding is negligible. The strong linear fit between the two variance estimates in both years suggests that equation 24 may be a reasonable way to specify the lower-tail variance estimate.

In general, β can vary by year. When this is the case, parameter estimates will be inconsistent. However, we can analytically characterize the magnitude of the bias. It is easy to show that the theoretical autocovariance matrix of the lower-tail estimates will have the same form as M , except that the parameters ψ_t will be replaced by $\psi'_t = \psi_t \sqrt{\frac{\beta_t}{\beta_1}}$. $\hat{\psi}_t$ will be biased, but we know that the bias factor is exactly the square root of the ratio $\frac{\beta_t}{\beta_1}$. A simple split-sample estimator of β_{79} and β_{89} (similar to that used in Tables 5A and 5B) yield estimates of 1.39 and 1.86, respectively. This implies a bias factor of about 1.157.

Table 6 reports the minimum distance estimates for both the between- and within-cohort

³⁷Unfortunately, the unattractive aspect of the lower-tail estimator in this application is that, unlike the conventional unbiased variance estimator, its small-sample properties are not analytically tractable. In general, the sampling error will not have a zero mean, and since the bias may vary over the J groups, it may also be correlated with the true wage variance. Thus the analysis in this section must necessarily assume that the number of observations per cell is sufficiently large in these data that the parameter estimates are not significantly affected.

³⁸ $\sigma_{u,t}^2$ will include a term that is the average variance of that year's specification error, $\omega_{.,t}$.

data. Because the number of cells that can be used for estimation sharply declines when all of the years are used, we produce one set of estimates using 4 years of data, and one using all 7 years of data. Column (1) reports the estimates when using the data from the 4 years which are least affected by top-coding. The top part of the table reports estimates of ψ_{81} , ψ_{89} , and ψ_{91} . It shows that by the end of the 1980s, the return to unobserved skill rose by about 30 percent, relative to 1979. Assuming the estimated bias factor of 1.157 is the same in 1989 and 1991, this would imply estimates of 1.069 and 1.130, respectively. The resulting goodness of fit statistic, asymptotically distributed as $\chi^2(2)$, is 4.06. The estimates of the nuisance parameters $Var[\sigma_{aj}^2]$ and $\sigma_{u,t}^2$ in the bottom part of Table 6 give an idea of how much of the variation in the within-group variances in each year is “signal”, and how much is “noise” due to sampling error. We see in Column (1) that for this particular specification, a little less than half of the observed variation in within group variances is “true” variation in within-group skill variances.

Column (2) reports the analogous estimates when all 7 years of data are used. The estimates for ψ_{89} and ψ_{91} are of similar magnitudes to those in Column (1). Note that the relative size of $Var[\sigma_{aj}^2]$ relative to $\sigma_{u,79}^2$ is somewhat smaller in this specification, suggesting that useful variation in variances was lost when the sample size declined from 237 to 140. The goodness of fit statistic, distributed as $\chi^2(14)$, is calculated as 24.8, suggesting that the data could reject our model in favor of a less restrictive one. However, the goodness of fit test in this particular specification should be viewed with caution, in light of the significantly smaller number of cells that results when using all 7 years of the data.

Columns (3) and (4) report the estimates from the within-cohort analysis. In many

respects, the within-cohort data yield very similar estimates of the change in the return to unobserved skill over the course of the 1980s. The one exception is that the implied return to unobserved skill declines relative to the base year, in the early part of the 1980s, before rising sharply in the late 1980s.

6 Implications for the College Premium and Discrimination

What do our estimates of the increase in the payoff to unobservable skill imply about the magnitudes of bias in conventional measures of the change in the college premium or conventional measures of changes in discrimination? Many of the alternative estimates of ψ_{89} or ψ_{91} presented in this analysis seem to be of a similar magnitude — around 1.10 to 1.20 for both the between- and within-cohort data. Given a 10 percent or a 20 percent rise in the return to unobserved skill, the appropriate columns ($\psi_{91} = 1.09$ or $\psi_{91} = 1.17$) of Table 1 gives an idea of how sensitive estimates of true changes in the college premium or discrimination would be to a wide range of assumptions about λ . It is clear that given an estimate of $\psi_{91} = 1.17$, the choice of λ seems to make little difference. Even in the extreme assumption that $\lambda = 1$, the college premium still rises on the order of .15 log points and the black-white differential still falls slightly. In short, the estimated ψ_{91} is simply not large enough to support the claim that the rise in the college-high school differential can be completely explained by an increase in the return to unobserved skill during the 1980s.

It is difficult to reconcile this claim with the data, even when we consider one of our largest estimates of ψ_{91} . By way of summary, Table 7 focuses on potential estimates of 1979-1991 changes in the college premium and discrimination as a function of λ , given that $\psi_{91} = 1.30$. Even though most of our estimates suggest that ψ_{91} is about .1 to .2 smaller,

it is instructive to construct an upper bound on how returns to unobserved skill could affect educational or race wage differentials.³⁹ The table uses the reduced-form wage differentials from Figures 2 through 4.

Table 7 shows that across the various experience cohorts, whether we examine between- or within-cohort movements, the college premium still rises significantly, within a wide range of beliefs about λ . Even under the extreme assumption of $\lambda = 1$ — that the return to college in 1979 is zero, the payoff to college rises between .07 and .17 log points for the youngest two cohorts, in the between- and within-cohort data.

Given this estimate of ψ_{91} , estimates of changes in discrimination are somewhat more sensitive to assumptions about λ . For example, in the 16-20 experience cohort, we see that under the assumption that the initial period's black-white wage gap was fully attributable to discrimination, we see that discrimination coefficient becomes more negative, by the end of the 1980s, by about .02 log points. On the other hand, if you assume that the initial period's wage gap is entirely due to an unobserved skill gap, then the implied discrimination coefficient would be positive by the end of the decade, about .03 log points in 1991.

However, even this generous estimate of the growth in the payoff to unobserved skill is unable to explain the widening black-white wage gap for the youngest cohort of workers. Even with the additional assumption that the discrimination coefficient in 1979 is zero, black workers in this cohort lose about .03 log points relative to their white counterparts. This result is even more pronounced for the relatively well-educated, youngest cohort. As the final column in Table 7 shows, the discrimination coefficient widens by nearly .14 log points, even

³⁹In particular, $\psi_{91} - 1.30$ is in the neighborhood of the estimates from Table 6. We reiterate, though, that our calculation of the relation between the conventional variance estimate and the lower-tail estimate suggests that this may be biased upward by about .20.

after one makes the extreme assumption that none of the .06 log point advantage for white workers in 1979 is due to discrimination. Even if there were “reverse” discrimination, with a coefficient of *positive* .06 log points, the discrimination coefficient would still fall about .11 log points, implying a discrimination coefficient of negative .05 log points by the end of the decade. In fact, it is easy to calculate that if the claim was made that the discrimination coefficient is zero in 1991 (meaning the unobserved skill gap is the entire black-white wage differential of -.21), then this would imply a decline in the discrimination coefficient of about .10 log points during the 1979-1991 period. This implies “reverse” discrimination — a .10 log point advantage for black workers in 1979.

In conclusion, our analysis of the data on within-group wage variances suggests that the rise in the return to unobservable skill is not large enough to be the sole explainer of the significant growth in the college-high school wage differential. However, we hesitate to further conclude that the “residual growth” in the college-high school wage gap (the growth that remains after parsing out the effects of a rising payoff to unobserved skill) is fully attributable to the relative market price of college-educated labor. In our econometric framework, such an inference is crucially dependent on the identifying restriction that $a_{COL} - a_{HS}$ is constant over time. Significant changes in the “quality” differences, or unobserved skill gap ($a_{COL} - a_{HS}$), between college and high school graduates over time, can have a first-order effect on the movements in the college-high school wage differential, independently of any change in the return to college. This is particularly the case in a between-cohort context.⁴⁰ Obtaining credible estimates of changes in relative worker quality (between

⁴⁰We argue that it is less of an issue in the within-cohort analysis. But, as noted earlier, in the presence of significant experience-education interactions, the interpretation of within-cohort changes in wage differentials is less clear. The information in longitudinal data cannot help address how the return to college for incoming

different educational groups) over time would better our understanding of the driving forces behind increased wage inequality.

Finally, we conclude that the hypothesis that the increasing wage gap between young black and white workers during the 1980s was largely due to a rise in the return to skill is largely inconsistent with even our largest estimates of the growth in the return to skill, and the most extreme assumptions about an unobserved skill bias between black and white workers.

labor market entrants changes over time.

References

- [1] Blackburn, McKinley, David Bloom, and Richard B. Freeman, "The Declining Position of Less-Skilled American Males," in *A Future of Lousy Jobs?* Gary Burtless, ed. (Washington, DC: The Brookings Institution, 1990), pp.31-67.
- [2] Blackburn, McKinley, and David Neumark, "Omitted-Ability Bias and the Increase in the Return to Schooling," *Journal of Labor Economics*, Vol. 11 (1993), No.3, 521-544.
- [3] Bound, John and Richard B. Freeman, "Black Economic Progress: Erosion of the Post-1965 gains in the 1980s?" in *The Question of Discrimination* Steven Shulman and William Darity, Jr., eds., (Middletown, CT: Wesleyan University Press, 1989), pp. 32-49.
- [4] Bound, John and Richard B. Freeman, "What Went Wrong? The Erosion of Relative Earnings and Employment Among Young Black Men in the 1980s," *Quarterly Journal of Economics*, CVII, No. 1 (February), 201-32.
- [5] Bound, John, and George Johnson, "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations," *American Economic Review*, June 1992, 371-392.
- [6] Card, David, "Earnings, Schooling, and Ability Revisited," *Industrial Relations Section Working Paper #331*, May 1994.
- [7] Card, David, and Thomas Lemieux, "Wage Dispersion, Returns to skill, and Black-White Wage Differentials," *NBER Working Paper No. 4365*, May 1993.
- [8] Card, David, and Thomas Lemieux, "Changing Wage Structure and Black-White Wage Differentials: A Longitudinal Analysis," *Industrial Relations Section Working Paper #322*, January 1994.
- [9] Chay, Kenneth Y., "Evaluating the Impact of the 1964 Civil Rights Act on the Economic Status of Black Men using Censored Longitudinal Earnings Data," mimeo, Princeton University, October 1995.
- [10] Deaton, Angus, "Panel Data from Time Series of Cross-Sections," *Journal of Econometrics*, Vol. 30, pp. 109-126.
- [11] Grogger, Jeff, "does School Quality Explain the Recent Black/White Wage Trend?" *Journal of Labor Economics*, Vol. 14, No. 2, pp. 231-253.
- [12] Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce, "Accounting for the Slowdown in Black-White Wage Convergence," in *Workers and Their Wages*, M. Kosters, ed. (Washington, DC: The AEI Press, 1991), pp. 107-143.
- [13] Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce, "Wage Inequality and the Rise in the Returns to Skill," *Journal of Political Economy*, Vol. 101, No. 3, pp. 410-442.
- [14] Katz, Lawrence F., and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, CVII (1992), 35-78.
- [15] Murphy, Kevin M., and Finis Welch, "The Structure of Wages," *Quarterly Journal of Economics*, CVII (1992), 285-326.

- [16] Taber, Christopher R., "The Rising College Premium in the Eighties: Return to College or Return to Ability," mimeo, Northwestern University, April 1996.
- [17] Willis, R. J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in *Handbook of Labor Economics*, Ashenfelter and Layard, eds.

Figure 1
Changes in Relative Wages in the 1980s: Educational, Black-White
Wage Differentials, and Within-group Wage Dispersion

Note: Quantities on the vertical axis denote the change from the base year (1979) in the (regression adjusted) college-high school and black-white log(wage) differential as well as the change in the overall residual standard deviation of log(wage). The 1979 levels of each series are .297, -.1478, and .398 for the college-high school, black-white wage differentials and the residual standard deviation, respectively. Details on their computation are described in the text.

Figure 2: College-High School Wage Differentials, 1979-1991, by experience cohort

Figure 3: Black-White Wage Differentials, 1979-1991, by experience cohort

Figure 4: Black-White Wage Differentials, 1979-1991, by education, by experience cohort

Figure 5: Residual Wage Dispersion (Standard Deviation), 1979-1991, by experience cohort

Figure 6: Fraction of college Graduates Top-Coded, 1979-1991

A. All Men and College Grads by Race

B. College Grads by Experience Cohort

**Figure 7A: Within-group Wage Variances,
Between-cohort: 1989 vs. 1979**

**Figure 7B: Within-group Wage Variances,
Within-cohort: 1989 vs. 1979**

**Figure 8A: Estimates of cell wage variances, 1979:
"Lower tail" vs. conventional variance estimate**

**Figure 8B: Estimates of cell wage variances, 1989:
"Lower tail" vs. conventional variance estimate**

Table 1

Potential Estimates of the Changes in the College Premium and Discrimination, 1979-1991

Constant "noise" variance over time								
Change in College Premium					Change in Discrimination			
$\theta = 0$ $\theta = .5$ $\theta = .9$					$\theta = 0$ $\theta = .5$ $\theta = .9$			
$\psi_{91} = 1$	$\psi_{91} = 1.09$	$\psi_{91} = 1.17$	$\psi_{91} = 1.67$		$\psi_{91} = 1$	$\psi_{91} = 1.09$	$\psi_{91} = 1.17$	$\psi_{91} = 1.67$
$\lambda = 0$	0.201	0.201	0.201	0.201	-0.0405	-0.0405	-0.0405	-0.0405
$\lambda = .25$	0.201	0.194	0.188	0.151	-0.0405	-0.0371	-0.0342	-0.0157
$\lambda = .5$	0.201	0.188	0.176	0.102	-0.0405	-0.0338	-0.0279	0.0090
$\lambda = 1$	0.201	0.174	0.151	0.002	-0.0405	-0.0272	-0.0154	0.0585

Declining "noise" variance. Noise variance (1991) = .9 * (Noise Variance (1979))								
Change in College Premium					Change in Discrimination			
$\theta = 0$ $\theta = .5$ $\theta = .9$					$\theta = 0$ $\theta = .5$ $\theta = .9$			
$\psi_{91} = 1$	$\psi_{91} = 1.09$	$\psi_{91} = 1.47$	$\psi_{91} = 1.92$		$\psi_{91} = 1$	$\psi_{91} = 1.09$	$\psi_{91} = 1.47$	$\psi_{91} = 1.92$
$\lambda = 0$	0.201	0.201	0.201	0.201	-0.0405	-0.0405	-0.0405	-0.0405
$\lambda = .25$	0.201	-	0.166	0.133	-0.0405	-	-0.0231	-0.0065
$\lambda = .5$	0.201	-	0.131	0.064	-0.0405	-	-0.0058	0.0275
$\lambda = 1$	0.201	-	0.062	-0.072	-0.0405	-	0.0290	0.0955

Note: Numbers are computed using data from Figure 1. λ refers to the fraction of the base year wage gap attributed to unobserved skill, θ refers to the fraction of the base year wage variance attributed to "noise", and ψ_{91} is the return to unobserved skill in 1991, relative to 1979. Details in text.

Table 2: Educational Distributions 1979-1991: Between- and Within-cohort changes
Male earners, by race, by experience cohort

Experience in 1979	Educ.	White Males			Black Males		
		1979 (Percent)	1991 Between (Percent)	1991 Within (Percent)	1979 (Percent)	1991 Between (Percent)	1991 Within (Percent)
1-5 Yrs.	<12	13.15	11.28	11.27	22.83	15.02	10.86
	12	41.64	35.92	39.86	43.89	46.57	44.92
	13-15	22.44	26.65	20.83	21.48	25.03	24.93
	16+	22.77	26.15	28.04	11.79	13.38	19.29
# of Obs.		14414	10272	12177	1187	919	1151
6-10 Yrs.	<12	11.13	10.33	10.19	21.58	8.59	14.13
	12	36.94	38.57	35.44	42.47	50.25	47.43
	13-15	22.5	21.83	22.44	21.49	24.18	22.33
	16+	29.42	29.26	31.92	14.47	16.98	16.11
# of Obs.		13531	11331	11099	1168	1001	1012
11-15 Yrs.	<12	13.02	11.58	11.48	22.82	10.1	17.83
	12	38.92	40.61	37.33	43.31	46.26	47.67
	13-15	22.4	20.14	23.48	21.13	24.08	21.83
	16+	25.66	27.68	27.71	12.74	19.57	12.66
# of Obs.		10195	12371	8144	942	1109	774
16-20 Yrs.	<12	16.91	10.47	14.53	32.48	12.75	27.89
	12	41.55	37.05	41.94	41.64	44.86	44.06
	13-15	18.49	21.9	19.35	15.9	23.94	17.59
	16+	23.06	30.59	24.17	9.97	18.44	10.46
# of Obs.		8044	11564	6263	742	1090	631
21-25 Yrs.	<12	21.55	10.45	19.85	39.88	16.16	37.01
	12	40.4	35.18	40.64	38.96	49.38	39.09
	13-15	15.71	24.14	17.46	11.72	21.58	14.55
	16+	22.34	30.23	22.05	9.44	12.88	9.36
# of Obs.		7117	9523	5047	657	885	481

Note: Computed from the CPS Merged Outgoing Rotation Group files. Above distributions are for those with nonmissing log(wages). Wage is the edited hourly earnings for workers paid by the hour, and for the workers not paid by the hour, computed by dividing the edited usual weekly earnings by edited usual weekly hours. Those with real wages (in 1991 dollars) less than \$2 or more than \$60 are dropped.

Table 3: Within-group wage variances and estimates of ψ_{89} , between and within experience cohorts, 1979-1989

			Between			Within		
1979			1989			1989		
Exper.	HS	COL	HS	COL	ψ_{89}	HS	COL	ψ_{89}
6-25	0.148	0.194	0.179	0.227	1.024 (0.056)	0.188	0.245	1.109 (0.064)
6-10	0.148	0.175	0.169	0.221	1.386 (0.160)	0.181	0.220	1.205 (0.157)
11-15	0.151	0.201	0.175	0.238	1.112 (0.101)	0.200	0.231	0.780 (0.124)
16-20	0.145	0.214	0.181	0.220	0.754 (0.085)	0.183	0.285	1.216 (0.113)
21-25	0.149	0.207	0.200	0.233	0.756 (0.118)	0.191	0.292	1.315 (0.141)
Total # of Obs.	16653	10400	19003	13230		13658	8819	
			Between			Within		
1979			1989			1989		
Exper.	Black	White	Black	White	ψ_{89}	Black	White	ψ_{89}
6-25	0.154	0.161	0.170	0.195	2.019 (0.692)	0.173	0.209	2.365 (0.793)
6-10	0.153	0.155	0.156	0.188	3.844 (6.282)	0.176	0.197	4.307 (13.273)
11-15	0.157	0.162	0.177	0.193	1.784 (1.497)	0.172	0.207	2.001 (0.965)
16-20	0.169	0.164	0.171	0.197	-	0.176	0.218	-
21-25	0.136	0.166	0.179	0.207	0.959 (0.224)	0.166	0.225	1.444 (0.257)
Total # of Obs.	3155	38634	3823	43405		2689	31107	

Note: Within-group variances are the estimated residual variance of a log(wage) regression on fully interacted sets of single-year experience, single-year education, and race dummies. The upper panel includes only those with 12 or 16+ yrs. of education. The lower panel includes all educational groups. Standard errors of the between-cohort and within-cohort estimates of ψ_{89} are in parentheses.

Table 4A: Summary statistics, cell-level data for between-cohort analysis

Variable		# of Cells	Mean	Std. Dev.	Min	Max
Age		503	42.2	11.7	20	64
Yrs. of Ed		503	12.9	3.1	6	18
Race		503	1.1	0.3	1	2
# of Obs.	1979	380	159.4	191.4	30	1150
	1981	363	170.0	210.0	30	1217
	1983	353	156.8	193.1	30	1087
	1985	337	164.8	220.2	30	1209
	1987	320	165.8	226.9	30	1096
	1989	410	165.0	203.3	30	1066
	1991	329	182.0	205.1	30	1024
Variance	1979	380	0.173	0.042	0.078	0.376
	1981	363	0.175	0.039	0.077	0.316
	1983	353	0.185	0.040	0.083	0.378
	1985	337	0.184	0.038	0.097	0.328
	1987	320	0.186	0.039	0.090	0.315
	1989	410	0.208	0.056	0.067	0.445
	1991	329	0.212	0.055	0.075	0.428
Variance (estimate #2)	1979	380	0.208	0.075	0.063	0.721
	1981	363	0.215	0.069	0.037	0.556
	1983	353	0.234	0.077	0.086	0.546
	1985	337	0.230	0.077	0.060	0.568
	1987	320	0.229	0.077	0.067	0.473
	1989	410	0.250	0.106	0.066	0.959
	1991	329	0.256	0.105	0.052	0.667

Note: Sample excludes cells for which there are fewer than 30 observations, more than 5 percent reporting wages less than the minimum wage, or at least one top-coded observation below the median, for each year.

Table 4B: Summary statistics, cell-level data for within-cohort analysis

Variable		# of Cells	Mean	Std. Dev.	Min	Max
Age		440	42.8	10.9	21	64
Yrs. of Ed		440	12.6	3.1	6	18
Race		440	1.1	0.3	1	2
# of Obs.	1979	363	154.5	186.1	30	1126
	1981	340	160.8	198.9	30	1217
	1983	313	140.5	177.5	30	990
	1985	271	133.8	179.1	30	1056
	1987	231	137.2	183.4	30	1017
	1989	277	133.4	161.3	30	964
	1991	209	152.7	169.3	30	833
Variance	1979	363	0.173	0.043	0.078	0.376
	1981	340	0.174	0.039	0.077	0.316
	1983	313	0.183	0.040	0.083	0.378
	1985	271	0.183	0.040	0.097	0.328
	1987	231	0.188	0.041	0.090	0.315
	1989	277	0.216	0.058	0.067	0.445
	1991	209	0.220	0.060	0.098	0.428
Variance (estimate #2)	1979	363	0.210	0.076	0.063	0.721
	1981	340	0.214	0.069	0.037	0.556
	1983	313	0.233	0.079	0.086	0.546
	1985	271	0.232	0.080	0.060	0.568
	1987	231	0.237	0.074	0.067	0.473
	1989	277	0.267	0.113	0.066	0.959
	1991	209	0.275	0.113	0.081	0.667

Note: Sample excludes cells for which there are fewer than 30 observations, more than 5 percent reporting wages less than the minimum wage, or at least one top-coded observation below the median, for each year.

Table 5A: Estimates of Changes in the Return to Unobserved Skill,
Between Cohort, 1979-1989: reduced form "slope" estimates, and implied ψ_{89}

	Slope (1)	1/Slope (2)	Ψ_{89} (3)	Obs	
Split-sample Estimate	1.121 (0.132)	0.746 (0.084)	1.110 (0.047)	1324	
Instrument:	Usual var. estimate		"Lower tail" estimate		
Year used as IV	Slope (1)	Ψ_{89} (2)	Slope (3)	Ψ_{89} (4)	
1981	1.189 (0.132)	1.090 (0.060)	1.351 (0.182)	1.162 (0.078)	274
1983	1.247 (0.167)	1.117 (0.075)	1.206 (0.173)	1.098 (0.079)	272
1985	1.823 (0.478)	1.350 (0.177)	1.684 (0.327)	1.298 (0.126)	253
1987	1.603 (0.403)	1.266 (0.159)	1.710 (0.361)	1.308 (0.138)	231
1991	1.226 (0.131)	1.107 (0.059)	1.213 (0.134)	1.101 (0.061)	269
(2SLS) 1981, 1991	1.187 (0.122)	1.089 (0.056)	1.250 (0.137)	1.118 (0.061)	237

Note: Robust standard errors in parentheses. Reduced form slope coefficients are from IV regressions (weighted) of the cell variances in 1989 on those of 1979, using other years' variances as instruments. Column (1) reports estimates using the sample variance estimates of the censored data, while column (2) are from using "lower tail" estimates of the within-group variance. Upper panel are estimates from a split-sample estimate, which is described in more detail in the text.

Table 5B: Estimates of Changes in the Return to Unobserved Skill,
Within Cohort, 1979-1989: reduced form "slope" estimates, and implied ψ_{89}

	Slope (1)	1/Slope (2)	Ψ_{89} (3)	Obs	
Split-sample Estimate	1.413 (0.197)	0.698 (0.083)	1.194 (0.058)	952	
Instrument:	Usual var. estimate		"Lower tail" estimate		
Year used as IV	Slope (1)	Ψ_{89} (2)	Slope (3)	Ψ_{89} (4)	
1981	1.262 (0.229)	1.124 (0.102)	1.493 (0.260)	1.222 (0.106)	207
1983	1.513 (0.377)	1.230 (0.153)	1.338 (0.316)	1.157 (0.136)	196
1985	1.244 (0.847)	1.115 (0.380)	1.716 (0.638)	1.310 (0.244)	175
1987	4.031 (1.928)	2.008 (0.480)	3.189 (1.078)	1.786 (0.302)	160
1991	1.577 (0.186)	1.256 (0.074)	1.417 (0.156)	1.190 (0.065)	180
(2SLS) 1981, 1991	1.360 (0.186)	1.166 (0.080)	1.431 (0.163)	1.196 (0.068)	163

Note: Robust standard errors in parentheses. Reduced form slope coefficients are from IV regressions (weighted) of the cell variances in 1989 on those of 1979, using other years' variances as instruments. Column (1) reports estimates using the sample variance estimates of the censored data, while column (2) are from using "lower tail" estimates of the within-group variance. Upper panel are estimates from a split-sample estimate, which is described in more detail in the text.

Table 6: Minimum Distance estimates of ψ_t , Between- and Within-cohort

		Between-cohort		Within-cohort	
		(1)	(2)	(3)	(4)
Return to Skill (Base=1979)	1979	1	1	1	1
		-	-	-	-
	1981	0.969 (0.070)	1.132 (0.172)	0.794 (0.075)	0.822 (0.174)
	1983	-	1.121 (0.159)	-	0.668 (0.226)
	1985	-	1.007 (0.146)	-	0.887 (0.195)
	1987	-	1.259 (0.158)	-	1.125 (0.170)
	1989	1.237 (0.083)	1.262 (0.171)	1.233 (0.070)	1.238 (0.173)
	1991	1.308 (0.092)	1.361 (0.162)	1.322 (0.088)	1.296 (0.159)
Skill variance (1979)		0.00109 (0.00025)	0.00033 (0.00014)	0.00139 (0.00032)	0.00033 (0.00015)
Sampling error var.	1979	0.00197 (0.00023)	0.00124 (0.00017)	0.00103 (0.00023)	0.00073 (0.00015)
	1981	0.00154 (0.00017)	0.00115 (0.00017)	0.00140 (0.00015)	0.00091 (0.00014)
	1983	-	0.00176 (0.00027)	-	0.00163 (0.00027)
	1985	-	0.00138 (0.00017)	-	0.00145 (0.00023)
	1987	-	0.00154 (0.00022)	-	0.00136 (0.00024)
	1989	0.00191 (0.00043)	0.00122 (0.00015)	0.00193 (0.00034)	0.00201 (0.00038)
	1991	0.00172 (0.00029)	0.00120 (0.00022)	0.00239 (0.00032)	0.00121 (0.00027)
Goodness of fit		4.062355	24.844	0.3585	71.0945
Degrees of freedom		2	14	2	14
Number of cells		237	140	163	92

Note: Estimated asymptotic standard errors in parentheses

Table 7

Estimates of the Changes in the College Premium and Discrimination, 1979-1991

$\psi_{91} = 1.30$

Change in College Premium (1979-1991)

	Between-cohort				Within-cohort			
	Years of exp. in 1979				Years of exp. in 1979			
	6-10	11-15	16-20	21-25	6-10	11-15	16-20	21-25
$\lambda = 0$	0.241	0.174	0.125	0.122	0.174	0.152	0.121	0.100
$\lambda = .25$	0.224	0.153	0.104	0.099	0.157	0.131	0.100	0.078
$\lambda = .5$	0.207	0.132	0.083	0.077	0.140	0.111	0.078	0.055
$\lambda = 1$	0.174	0.090	0.040	0.032	0.106	0.069	0.036	0.010

Change in Discrimination (1979-1991)

	Within-cohort				Within cohort (6-10 yrs. of exp in 1979)	
	Years of exp. in 1979					
	6-10	11-15	16-20	21-25	Ed $\leq$ 12	Ed $\geq$ 13
$\lambda = 0$	-0.069	0.028	-0.019	-0.006	-0.059	-0.155
$\lambda = .25$	-0.060	0.041	-0.007	0.006	-0.047	-0.150
$\lambda = .5$	-0.051	0.054	0.005	0.018	-0.036	-0.146
$\lambda = 1$	-0.032	0.079	0.030	0.042	-0.012	-0.137

Note: Numbers are computed using data from Figures 2-5. λ refers to the fraction of the base year wage gap attributed to unobserved skill.