

The Longevity Gains of Education

Yuri Sanchez

7/31/2010

This paper analyzes the longevity benefits of education in the form of improved life expectancy, monetized survivorship gains and longevity-adjusted Mincerian returns. Mortality data from the NLMS reveal a negative and statistically significant effect of education on eleven-year mortality in the order of -0.005 (OLS and Cox PH), -0.016 (IV-2SLS) and -0.021 (IV Cox PH). Region of birth and US compulsory schooling laws in place between 1925 and 1939 serve as instruments for schooling. These marginal effects correspond to improvements in life expectancy at age 25 of 0.28, 0.34 and 0.40 years, respectively. Using a lifetime expected utility maximization model to monetize longevity gains yields the following longevity-adjusted Mincerian returns to schooling: 7.1% for white males (6.2% income, 0.9% longevity), 7.0% for white females (6.5% income, 0.5% longevity), 5.6% for black males (5.3% income, 0.3% longevity), and 4.9% for black males (4.6% income, 0.3% longevity). Longevity gains accrue well into late-adulthood, crest around age 50 and decrease with specialization in proportion to income gains. IPUMS data reveal that the trends of the marginal effect of education on income and mortality present marked similarities across race and gender groups between 1940 and 1990. Longevity returns are rapidly rising since 1970 but this tendency reverses for black males in 1980. Steadily increasing since 1950, Mincerian returns soar after 1980, especially among blacks. Longevity and income estimates are remarkably similar across NLMS and IPUMS datasets in spite of different mortality and income measures, providing consistency and robustness to the results.

Contents

Tables and Figures	3
1. Introduction.....	4
2. A Model to Measure the Longevity Gains of Education	5
3. The Impact of Education on Mortality.....	8
3.1 NLMS Data Description.....	8
3.2 Empirical Estimation.....	12
3.2.1 The Education-Mortality Gradient.....	12
3.2.2 OLS and Cox PH Models	13
3.2.3 IV-2SLS and IV Cox PH Models	15
4. Longevity Gains of Education.....	23
4.1 Life Expectancy and Monetized Longevity Gains	23
4.2 Income Gains.....	26
4.3 Longevity-Adjusted Mincerian Returns.....	28
5. Evolution of Longevity and Income Returns to Education: 1940-1990	30
5.1 IPUMS Data Description	30
5.2 Longevity and Income Returns to Education: 1940-1990.....	32
6. Discussion and Conclusions	36
Bibliography.....	39
Datasets	40
Appendix	41
A1. Contrasting methodology and results with Lleras-Muney (2004)	41
A2. Measurement error bias from selective attrition and quality of education	42
A2.1 NLMS.....	42
A2.2 IPUMS	46
A3. Measurement error and aggregation biases with synthetic-cohort mortality measures.....	47
Tables	50
Figures	62

Tables and Figures

Table 1. Calibration of parameters.....	50
Table 2. NLMS sample restrictions.....	51
Table 3. NLMS summary statistics with missing and not missing state of birth (SOB) characteristics	52
Table 4. The schooling–mortality gradient: OLS and Cox PH models	53
Table 5. IV-2SLS weak instruments and over-identification tests.....	54
Table 6. The IV 1 st Stage: the effect of instruments on schooling	56
Table 7. Effect of schooling on 11-year mortality: OLS, Cox PH, IV-2SLS and IV Cox PH models.....	57
Table 8. Life expectancy and discounted longevity gains of schooling	59
Table 9. Mincerian returns to education by gender and race: OLS & IV-2SLS.....	60
Figure 1. The value of a life year $v(t)$ over the lifecycle	62
Figure 2. Eleven-year death rates in 1983 by age: NLMS <i>vs.</i> HMD	63
Figure 3. Log differences in eleven-year death rates by age: NLMS <i>vs.</i> HMD.....	64
Figure 4. Life expectancy at age 25 by years of schooling.....	65
Figure 5. Change in life expectancy (at selected ages) by years of schooling	66
Figure 6. Percentage (ψ_1) and level effects ($\psi_1\lambda$) of Cox PH regressions by age group	67
Figure 7. Translating CSL into CYS: the case of California.	68
Figure 8. Survivorship gains (top) and monetized longevity gains of schooling (bottom): Cox PH and IV Cox PH models	69
Figure 9. Present value of total (longevity plus income) gains of an additional year of schooling ($r = 3\%$) ..	70
Figure 10. Total (longevity θ plus income R) returns to schooling ($r = 3\%$).....	71
Figure 11. Description of IPUMS death rates	72
Figure 12. Comparison of age-standardized death rates: IPUMS <i>vs</i> HMD.....	73
Figure 13. OLS and IV-2SLS estimates of β_I : 1940-1990	74
Figure 14. IV-2SLS estimates ($\hat{\alpha}_1^{IV}$) by gender and race: 1940-1990.....	75
Figure 15. OLS Mincerian returns to schooling by gender and race: 1940-2000	76

1. Introduction¹

Schooling is a fundamental factor underlying the accumulation of human capital and eventual economic growth. Although most of the economic literature analyzing the link between schooling and social outcomes originally focused on income, a growing vein of research has shifted attention to other non-pecuniary outcomes like health² (morbidity, hygiene, diet, addictions and exercise); criminal behavior;³ marriage and divorce; teen pregnancy;⁴ migration; voting patterns and political participation.⁵ These studies have at heart revealed a subtle but important distinction between two concepts that may have previously seemed identical: education is in essence the human capital content of completed years of schooling. In addition, the study of non-pecuniary benefits has broadened the notion of formal schooling from only an agent of economic growth to a determinant of human development.

Joining this strand of the literature, the present study aims to shed light on the longevity gains of education focusing on three particular goals: i) identify the causal impact of education on mortality; ii) to express this effect in terms of survivorship benefits, life expectancy improvements, present value monetized gains, and longevity-adjusted Mincerian rates of return;⁶ and iii) to compare long run trends of income and longevity returns from 1940 to 1990. To assess education and mortality disparities across demographic groups, these goals are analyzed separately by race and gender.

The paper is organized as follows: Section 2 presents the model to measure the longevity gains of schooling in dollar terms. Section 3 identifies the causal impact of education on mortality using data from US compulsory schooling laws and the National Longitudinal Mortality Study. Section 4 presents estimates of the resulting survivorship benefits, life expectancy improvements, present value monetized gains, and longevity-adjusted Mincerian rates of return. The trends of longevity and income returns from 1940 to 1990 based on IPUMS data are presented in Section 5. Finally, Section 6 summarizes the paper's most relevant findings and highlights future areas of research.

¹ I would like to thank Gary Becker, Hoyt Bleakley, David Meltzer, and Kevin Murphy for their valuable guidance and encouragement. I am also grateful to Robert Fogel, Steven Levitt, Derek Neal, Hugo Sonnenschein, and Micro Lunch Workshop participants at the University of Chicago for providing helpful comments. Financial support by CONACYT Grant No. 178971 is acknowledged. This paper uses data supplied by the National Heart, Lung, and Blood Institute, NIH, DHHS from the National Longitudinal Mortality Study. The views expressed in this paper as well as any errors are my own and do not necessarily reflect the views of the National Heart, Lung, and Blood Institute, the Bureau of the Census, or the National Center for Health Statistics. Please address comments to: yuri@uchicago.edu.

² (Grossman, The Correlation between Health and Schooling, 1975) Grossman (1975).

³ (Lochner & Moretti, 2004). Lochner and Moretti (2004).

⁴ (Heckman, Stixrud, & Urzua, 2006). Heckman, Stixrud and Urzua (2006)

⁵ (Grossman, 2003). Grossman (2003).

⁶ Also known as private income returns, named after the pioneering work on the subject by Jacob Mincer (1974).

2. A Model to Measure the Longevity Gains of Education

The effect of education on mortality can be monetized using the framework developed by Murphy & Topel (2006). Consider the following lifetime expected utility maximization problem of a representative agent at age a :

$$\begin{aligned} & \text{Max}_{\{c(t), l(t), E\}} U(a) = \\ & = \int_a^\infty \{H(t)u(c(t), l(t))e^{-\rho(t-a)} + \mu[y(t; E) - c(t)]e^{-r(t-a)}\} \tilde{S}(t, a; E) dt + \mu[A(a) - K(E)] \end{aligned} \quad (1)$$

Ignoring the effects of education on morbidity, $H(t)$ is the stock of health at age t which does not depend on years of schooling E . Utility at age t , $u(c(t), l(t))$, is an increasing and concave function of consumption $c(t)$ and leisure $l(t)$. The agent discounts utility intertemporally at the rate ρ , and the market discounts resources at the interest rate r . Total income is assumed to depend only on labor earnings at every age: $y(t; E) = w(t; E)(1 - l(t))$. The wage rate $w(t; E)$ depends positively on education E and the total time endowment is normalized to unity. The Lagrange multiplier μ captures the marginal utility of lifetime resources in present value. The survival probability from age a to age t is given by $\tilde{S}(t, a; E) \equiv \prod_{\tau=a}^t (1 - \lambda_1(\tau; E)) \cong e^{-\int_a^t \lambda_1(\tau; E) d\tau}$, where $\lambda_1(t; E)$ is the annual hazard rate, or one-year probability of death conditional on reaching age t . Finally, $A(a)$ denotes initial asset holdings at age a and $K(E)$ is the up-front monetary cost of reaching schooling level E . In this model, education is an once-in-a-lifetime investment made by age a that pays future longevity and earnings returns. This optimization problem yields three first-order conditions:

$$H(t)u'_c(c(t), l(t)) = \mu e^{-(r-\rho)(t-a)} \quad (2)$$

$$H(t)u'_l(c(t), l(t)) = \mu w(t; E) e^{-(r-\rho)(t-a)} \quad (3)$$

$$\begin{aligned} & \int_a^\infty \left[\frac{\partial \tilde{S}(t, a; E)}{\partial E} \{H(t)u(c(t), l(t))e^{-(\rho+r)(t-a)} + \right. \\ & \left. \mu[y(t; E) - c(t)]\} \mu \frac{\partial y(t; E)}{\partial E} \right] e^{-r(t-a)} \tilde{S}(t, a; E) dt = \mu K'(E) \end{aligned} \quad (4)$$

To simplify the expression capturing income and longevity returns to schooling, $u(c(t), l(t))$ is assumed homothetic as in Murphy and Topel (2006). Utility $u(c, l)$ can thus be expressed as $u(z(c, l))$, where the composite good $z(c, l) = z_c c + z_l l$ is a monotone and linear function of consumption and leisure. Combining this result with equations (2) and (3) yields a wage rate which only depends on the positive constants z_c and z_l :

$$w(t; E) = \frac{u'_l(c(t), l(t))}{u'_c(c(t), l(t))} = \frac{z_l}{z_c} \quad (5)$$

The total (longevity plus income) return to an additional year of schooling from education level E at age a is defined as the marginal rate of substitution between education and initial assets:

$$V_T(a; E) \equiv \frac{\partial U(a)/\partial E}{\partial U(a)/\partial A(a)} = \frac{\partial U(a)/\partial E}{\mu} = \int_a^\infty \left\{ v(t) \frac{\partial \tilde{S}(t, a; E)}{\partial E} + \frac{\partial y(t; E)}{\partial E} \right\} e^{-r(t-a)} \tilde{S}(t, a; E) dt - K'(E) \quad (6)$$

The second equality in equation (6) follows from the definition of μ . The third equality follows from the assumption of homotheticity which allows the value of life-year t ($v(t)$) to be expressed as a linear combination of full income ($y^F(t; E) \equiv y(t; E) + w(t; E)l(t)$) and full consumption ($c^F(t) \equiv c(t) + w(t; E)l(t) = \frac{z}{z_c}$):

$$v(t) \equiv \frac{u(c(t), l(t))}{u'_c(c(t), l(t))} + y(t; E) - c(t) = y^F(t; E) + c^F(t) \Phi(z) \quad (7)$$

The factor $\Phi(z) \equiv \left[\frac{u(z)}{zu'(z)} - 1 \right]$ is a measure of consumer surplus per unit of composite good z . The term is positive as long as average utility ($\frac{u(z)}{z}$) exceeds marginal utility ($u'(z)$). As $u(z)$ becomes more concave, the intertemporal elasticity of substitution falls and $\Phi(z)$ rises.

Ignoring morbidity effects ($\frac{\partial H}{\partial E} = 0$) and costs of schooling ($K'(E) = 0$), the present value discounted total (longevity plus income) gains from a marginal increase in schooling E at age a can be expressed as:

$$V_T(a; E) \equiv \int_a^\infty \left\{ v(t) \frac{\partial \tilde{S}(t, a; E)}{\partial E} + \frac{\partial y(t; E)}{\partial E} \right\} e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (8)$$

Equation (8) shows that an additional year of schooling induces, in every period t , a longevity gain $(v(t) \frac{\partial \tilde{S}(t, a; E)}{\partial E})$ and an income gain $(\frac{\partial y(t; E)}{\partial E})$. The lifetime sums of these benefits discounted by the factor $S(t; E) \equiv e^{-r(t-a)} \tilde{S}(t, a; E)$ determine the present value longevity $V_L(a; E)$ and income $V_Y(a; E)$ gains:

$$V_L(a; E) \equiv \int_a^\infty \left\{ v(t) \frac{\partial \tilde{S}(t, a; E)}{\partial E} \right\} e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (9)$$

$$V_Y(a; E) \equiv \int_a^\infty \left\{ \frac{\partial y(t; E)}{\partial E} \right\} e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (10)$$

Equation (8) is the basis to empirically quantify the total gains of education using mortality data from the NLMS.

Three additional assumptions are necessary to estimate the value of a life-year $v(t)$. First, net savings are zero throughout the lifecycle, so $y(t; E) = c(t)$ for all t . Second, agents are assumed to work 8 hours every day, which is one third of their total time endowment. Therefore, leisure is $l(t) = 2/3$, income equals one third of the wage rate for all t , and full income is equal to full consumption (and to the wage rate):

$$y^F(t; E) = c^F(t) = w(t; E) = 3y(t; E) \quad (11)$$

Third, the utility function takes the constant elasticity of substitution (CES) form

$$u(z) = \frac{z^{1-\sigma^{-1}} - z_0^{1-\sigma^{-1}}}{1-\sigma^{-1}} \quad (12)$$

where σ is the intertemporal elasticity of substitution (IES) and z_0 is the subsistence level such that $u(z_0) = 0$. This implies that consumer surplus per unit of composite good becomes:

$$\Phi(z) \equiv \left[\frac{u(z)}{zu'(z)} - 1 \right] = \frac{1}{\sigma-1} \left[1 - \sigma \left(\frac{z_0}{z} \right)^{1-\sigma^{-1}} \right] \quad (13)$$

With these three assumptions, the value of a life-year simplifies to:

$$v(t) = w(t; E)[1 + \Phi(z)] \quad (14)$$

Equation (8) shows that the intertemporal discount factor does not affect the valuation of longevity gains. This is the result of the perfect credit markets assumption embedded in the

intertemporal budget constraint. The interest rate is thus the only measure of opportunity cost for consuming and investing in any given period. To let individuals face different opportunity costs, 3%, 5%, 7% and 10% interest rates are employed in the estimation of longevity and income gains.

The statistical value of life (SVL) is computed as the sum of $v(t)$ from ages 25 to 90 discounted by $S(t; E)$. Viscusi and Aldy (2003) suggest a range for the SVL between \$5.5 and \$7.5 million and Murphy and Topel (2006) set the SVL at \$6.3 million. With an interest rate of 3% and a subsistence level $\frac{z_0}{z}$ of 0.10, consumer surplus $\Phi(z)$ is calibrated at 0.64 to yield an IES of $\sigma = 1.46$ and a SVL equal to \$5 million.⁷ Table 1 reports the calibration of these parameters as well as the average value of a life year, the value of a life year at age 50 ($v(50)$) and the SVL by gender and race. Higher interest rates cause a reduction in the SVL but, as equations (13) and (14) show, do not affect $\Phi(z)$, $v(t)$, or σ .

[Table 1]

Figure 1 shows the value of a life-year over the lifecycle by gender and race. This profile is concave, peaking between ages 45 and 55 and leveling off after age 75. At all ages, whites and males have a greater $v(t)$ relative to blacks and females, respectively.

[Figure 1]

3. The Impact of Education on Mortality

3.1 NLMS Data Description

The National Longitudinal Mortality Study (NLMS, Release 3) is a representative sample of the non-institutionalized US population on April 1, 1983. It is comprised of 11 CPS surveys matched to the National Death Index until 1994.⁸ The baseline survey in 1983 provides information on age, completed years of schooling, household income, race, gender, marital status, employment status, state of birth (SOB) and state of residence.

⁷ Murphy and Topel (2006) calibrate these parameters as follows: $\Phi(z) = 2.11$; $\sigma = 0.80$; $\frac{z_0}{z} = 0.10$; SVL=\$6.3 million.

⁸ Two-thirds of the sample are matched through the social security number. Mortality information is collected through the National Center for Health Statistics.

The primary advantage of the NLMS over mortality data from death certificates is that education and other individual characteristics are collected from a survey in a controlled and standardized setting. Information from death certificates is less detailed, often incomplete and prone to greater measurement error as third persons report it on behalf of the deceased.

People under 25 years of age, foreign-born individuals and natives or residents from Hawaii and Alaska are excluded from the sample to maximize the likelihood that schooling measures final completed years acquired in the US. These restrictions would reduce the initial sample size of 988,346 observations and 89,909 identified deaths to 545,557 individuals and 82,654 deaths. However, there is an additional source of missing data: 56% of the sample does not report SOB.⁹ The restricted sample with non-missing SOB has 242,276 observations and 38,348 identified deaths between 1983 and 1994.

[Table 2]

As Table 2 shows, 54% of casualties of US natives older than 24 years are lost with the exclusion of observations with missing SOB. Nonetheless, summary statistics reveal that this considerable source of attrition is fairly random. Table 3 compares the mean and standard deviation of eleven-year mortality, education, age and other socio-demographic characteristics across individuals whose SOB is observed to those whose SOB is missing. Mean and standard deviation differences are statistically significant across subsamples but small in magnitude. As main exceptions, non-black racial minorities (*Hispanic* and *other race* dummies) and the unemployed are over-represented in the subsample with missing SOB. In addition, OLS eleven-year mortality regressions carried out separately on these subsamples revealed a coefficient on education of -0.004 (standard error of 0.0002) from the subsample excluding SOB, and of -0.003 (standard error of 0.0003) in the subsample including SOB.¹⁰ The difference between these estimates is not statistically significant, suggesting attrition is sizable but not selective. As a result, estimates based on regressions including SOB have larger standard errors (from a lower sample size) but do not significantly suffer from

⁹ This attrition source is only present if the effect of education on mortality is estimated with IV-2SLS, since compulsory schooling laws, the instruments for education, are based on each individual's SOB.

¹⁰ The set of controls included in these regressions consists of: the natural logarithm of per capita income and dummies for: gender, state of residence, five-year age group, race, marital status, urban residence, employment status, and inter-state migration.

selection bias. Conversely, the omission of SOB in the regression models does not introduce a severe omitted variables problem.¹¹

[Table 3]

A potentially graver source of selection bias is the exclusion of institutionalized individuals. If this group has lower education and greater mortality relative to non-institutionalized people, death rates and the effect of education on mortality will be underestimated. To assess the magnitude of this problem, Figure 2 compares age-specific eleven-year mortality rates in 1983 to those from the Human Mortality Database (HMD). The HMD, unlike the NLMS, includes institutionalized individuals but does not provide mortality information by race, education level or any other characteristic aside from age and gender.¹² To make rates comparable, the HMD annual probability of death at age t ($\lambda_1(t)$) was converted to an eleven-year mortality rate $\lambda_{11}(t)$ using the formula: $\lambda_{11}(t) \cong 1 - e^{-\sum_{\tau=0}^{10} \lambda_1(t+\tau)}$.¹³

[Figure 2]

Relative to the HMD gender-specific rates, whites' NLMS mortality rates are nearly the same up to age 65, remain no more than ten percentage points above until age 85 and fall below by at most nine percentage points afterwards. Black mortality rates, on the other hand, are well above HMD rates before age 78 and fall considerably below afterwards. Disparities larger than ten percentage points occur after age 55 for black males and beyond age 85 for black females.

Figure 3 shows the log differences in mortality across genders and races relative to the NLMS average (top panel) and to HMD death rates (bottom panel). The top panel shows how initial mortality disparities relative to the sample average eventually decline with age. The convergence of death rates across demographic groups, or reversion to the mean, results from the pool of survivors

¹¹ The results from Table 4 were replicated without SOB and inter-state migration. In most cases the coefficient of education, our variable of interest, is not significantly affected although standard errors increase. This result is important since IV-2SLS estimates presented in Table 5 exclude SOB dummies as controls but include them as instruments interacted with compulsory years of schooling.

¹² The HMD collects individual death records from Mortality Detail Files and Multiple Cause of Death files disseminated by the National Center for Health Statistics.

¹³ The probability of surviving 11 years at age t under a variable annual hazard is $\tilde{S}_{11}(t) = \prod_{\tau=0}^{10} (1 - \lambda_1(t + \tau))$, which implies that $\ln(\tilde{S}_T) = \sum_{\tau=0}^{10} \ln(1 - \lambda_1(t + \tau)) \cong - \sum_{\tau=0}^{10} \lambda_1(t + \tau)$ and $\tilde{S}_{11}(t) \cong e^{-\sum_{\tau=0}^{10} \lambda_1(t+\tau)}$. Finally, note that $\lambda_{11}(t) = 1 - \tilde{S}_{11}(t) \cong 1 - e^{-\sum_{\tau=0}^{10} \lambda_1(t+\tau)}$.

at any given age having lower frailty (*i.e.* better health) than average. As age advances, the frailty distribution is progressively truncated from below reducing the dispersion of health and mortality. Blacks show reversion to the mean up to the age of 75 but subsequently have lower-than-average mortality by at most ten percentage points. This pattern suggests the exclusion of institutionalized individuals underestimates black mortality at elder ages.

[Figure 3]

From the bottom panel of Figure 3, the average NLMS death rate between the ages of 25 and 45 is 10% smaller than the corresponding HMD death rates. This effect is largely explained by the fact that incarceration accounts for a large fraction of institutionalized individuals at these ages, and inmates face greater risk of death than non-institutionalized individuals. This degree of underestimation gradually decreases and eventually becomes a slight overestimation between ages 60 and 80. Thus, mortality in the NLMS is underestimated for people younger than 60 by no more than 14% and overestimated for ages older than 60 by less than 8%. After age 80, health institutions account for most of the institutionalized population, resulting in the underestimation of mortality in the NLMS. On average, the NLMS overestimates black death rates between the ages of 25 and 60 by 15% for males and 20% for females. These levels decline subsequently to become an underestimation of no more than 13% past the age of 75. Since HMD mortality rates are not race-specific, this overestimation is to some extent natural as blacks have greater mortality than whites regardless of gender. By the same logic, slight underestimation white mortality is also expected. The top panel of Figure 3 provides some guidance on how large these dissimilarities should be.

Overall, evidence from Figures 2 and 3 suggest that the attenuation bias induced by the exclusion of institutionalized individuals mainly affects blacks at older ages, and should not be more than 13%. The underestimation of death rates at early ages by approximately 10% does not appear to be a serious source of selection bias as mortality rates under age 45 are low and level differences are quite small.

3.2 Empirical Estimation

3.2.1 The Education-Mortality Gradient

The relationship between mortality and education has two causal explanations. First, education could directly impact mortality through healthier and less risky habits, more or better health investments, or improved health knowledge. Second, education and mortality could be spuriously correlated through unobservable factors like the intertemporal discount rate;¹⁴ cognitive and non-cognitive skills;¹⁵ adverse experiences during childhood;¹⁶ frailty or poor health conditions;¹⁷ and parental inputs like income, education, time and genes. It may also be argued that high mortality risk could discourage education as the horizon over which to reap these investments is shortened. Instead of establishing a new causal link from mortality to education, this story is consistent with the claim that poor health early in the lifecycle causes less education and higher risk of death.

Schooling and life expectancy have a positive relationship regardless of gender and race, as Figure 4 shows. Life expectancy at age a for a person with E years of schooling is computed as:

$$LE(a; E) = \int_a^\infty \tilde{S}(t, a; E) dt \quad (15)$$

The survival probability from age a to age t is defined in terms of the annual hazard rate $\lambda_1(\tau; E)$ or one-year probability of death conditional on reaching age τ : $\tilde{S}(t, a; E) \equiv \prod_{\tau=a}^t (1 - \lambda_1(\tau; E))$.

[Figure 4]

With the exception of whites, schooling is positively related to life expectancy only after the 10th year. Moreover, the completion of high school (12th year) and college (16th year) appear to be associated with the greatest improvements in life expectancy. This feature is confirmed in Figure 5 by plotting the change in life expectancy at selected ages associated with an additional year of schooling: $\Delta LE(a; E) \equiv LE(a; E + 1) - LE(a; E)$. The completion of high school is associated with large improvements in life expectancy across gender and race: 2.5 years for black males, 2.2

¹⁴ Fuchs (1982).

¹⁵ Cunha and Heckman (2007)

¹⁶ Felitti (2002) and Peña (2009).

¹⁷ Frailty introduces survivorship bias. Less educated individuals who survive to older ages have lower frailty (*i.e.* are healthier) than average, so their death rate is similar to the average mortality of the surviving population which is more educated. Like the exclusion institutionalized individuals, survivorship-biased estimates underestimate the true effect of education on mortality.

years for black females, 1.5 years for white females and 1.3 years for white males. The completion of college is associated with large life expectancy improvements only in the cases of males (3.5 years for blacks and 1.8 years for whites). These positive correlations persist over time but gradually diminish with age.

[Figure 5]

The existence of “jumps” in life expectancy upon the completion of selected schooling levels suggests two non-mutually exclusive hypotheses. First, schooling could improve those aspects of human capital directly linked to longevity at these schooling levels more than at any others. In this sense, Figure 5 captures the “value added” associated to each schooling level. Second, underlying factors strongly associated with the completion of these schooling levels (*i.e.* cognitive and non-cognitive skills, intertemporal discount rates and parental inputs) could be the real determinants of improved longevity instead of the human capital acquired through schooling. The reason why such unobservables would strongly affect the completion of certain schooling years and not others should rely on greater psychic and monetary costs to complete these levels.

3.2.2 OLS and Cox PH Models

Bearing in mind the potential endogeneity of schooling and ignoring the non-linearity of its longevity returns, Table 4 presents the OLS estimates of the effect of schooling on eleven-year mortality (β_1) by gender, race and age group. These are obtained from the linear probability model:

$$D_i = \beta_0 + \beta_1 E_i + \beta_2' X_i + u_i \quad (\text{OLS})$$

The dependent variable D_i equals one if individual i died at most eleven years after being surveyed. Regressor E_i measures years of completed schooling and X_i denotes a set of controls including the natural logarithm of per capita household income¹⁸ and dummies for: SOB, state of residence, five-year age groups, marital status (single, widowed and divorced; married is the reference category), urban residence, inter-state migration¹⁹ and employment status (unemployed, disabled, out of the

¹⁸ Taking into account gains from specialization within the household, this variable is computed as the average household income divided by the square root of household size. All dollar measures based on NLMS data, including the monetized longevity gains of schooling, are expressed in 2004 USD.

¹⁹ This variable equals one if the person resides in a different state where he/she was born.

labor force; employed is the reference category). To address the problem of heteroskedasticity, White's robust standard errors are reported in parentheses.

OLS estimates reveal that for most gender, race and age groups there is a negative and statistically significant relationship between education and mortality. The gradient is steeper for whites relative to blacks, and likewise for males relative to females. These gender and racial differences, however, are not statistically significant. Estimates vary considerably across age groups; the level effect of education on mortality is greatest between ages 45 and 64. As suggested by Figures 2 and 3, OLS estimates for blacks older than 64 years could be attenuated from the exclusion of institutionalized individuals. This feature combined with a low sample size explains why estimates for blacks are small relative to whites and not statistically significant.

[Table 4]

Given the nature of NLMS data, the Cox proportional hazards (PH) duration model is more suitable for survival analysis:²⁰

$$\lambda_i(t = 11; E_i, X_i) = \theta_t e^{\psi_1 E_i + \psi_2' X_i} \quad (\text{Cox PH})$$

For each individual i in the sample, the variable $\lambda_i(t; E_i, X_i)$ captures the t -year hazard rate, where θ_t is the time-dependent hazard factor. OLS and Cox PH estimates are not directly comparable since β_1 captures a level effect while the semi-elasticity ψ_1 identifies a percentage change. However, they are approximately related by the formula: $\beta_1 \cong \psi_1 \lambda(t = 11; E, X)$.²¹ To facilitate comparisons across models, Table 4 reports the transformed Cox PH level effects $\psi_1 \lambda$.²²

The Cox PH quasi-maximum likelihood duration model improves efficiency by ranking casualties according to their day of death after survey. For example, a person who died one day after being interviewed in 1983 has a different contribution to the likelihood function than another who died 11 years later in 1994. OLS estimates, on the other hand, ignore this valuable information and treat both casualties equally.

Most of the results obtained with OLS are confirmed with the Cox PH model. The efficiency of all estimates is improved, especially in the black subsample. All estimates as well as the racial and

²⁰ In both OLS and Cox PH models, data is weighted by individual sampling weights.

²¹ Letting $E[D|E, X] \cong \lambda(t = 11; E, X)$, the formula follows from: $\psi_1 = \frac{\partial \lambda}{\partial E} \frac{1}{\lambda} = \frac{\beta_1}{\lambda}$.

²² Standard errors were not scaled by λ since they still yielded statistically significant estimates.

gender gaps become statistically significant. The marginal level effects become less important with age relative to the rapidly rising risk of death. An education-mortality gradient close to zero for black males ages 25 to 45 could be associated to the dramatic rise in incarceration rates of black males during the 1980s. Overall, Cox PH and OLS estimates are not very different across gender, race and age groups. Cox PH level estimates are smaller or equal to OLS estimates at ages younger than 65 and greater otherwise. With the exception of white males 65 years or older, level differences across models are not larger than two tenths of a percentage point.

To further explore the effect of schooling on mortality at the different stages of life, Figure 6 presents the percentage (ψ_1) and level effects ($\psi_1\lambda$) of Cox PH regressions based on 5-year age subsamples.²³ The decreases in mortality associated with an additional year of schooling are long-lived and greater for whites than blacks throughout the lifecycle. The greatest mortality reductions in percentage terms arise before age 45 and gradually decrease with age. Level reductions in the mortality, however, are very small early in the lifecycle but grow progressively until age 70 for whites and age 80 for blacks. This suggests that modest level reductions in mortality at early ages translate into large percentage changes. Although not reported, these patterns hold across genders within each race.

[Figure 6]

3.2.3 IV-2SLS and IV Cox PH Models

To consider estimates from Table 4 as causal effects of education on mortality it is necessary to eliminate the spurious correlation caused by unobservable factors (omitted variables). For this purpose, my IV approach exploits exogenous variation in education induced by state compulsory schooling laws (CSL).²⁴ Established in the second half of the 19th century, CSL suffered considerable modifications between 1915 and 1939 in 36 of the 51 US states affecting the schooling decision of individuals born between 1901 and 1932 in such treated states. Identification of the causal effect of education on mortality through IV is therefore possible only for individuals aged

²³ The set of controls includes the natural logarithm of per capita income and dummies for: female gender, state of residence, marital status, urban residence and employment status. Data is weighted by individual sampling weights and the dotted lines represent the estimates' 95% confidence intervals.

²⁴ Lleras-Muney (2004) provides a comprehensive description of the data available at: <http://www.econ.ucla.edu/alleras/research/data.html>.

between 51 and 82 years in 1983. Even though this group comprises only 36% of the sample, it accounts for almost 78% of identified deaths.²⁵

Changes in CSL affect education at the cohort and state-of-birth level. Hence, it is unlikely that these law amendments are correlated with individual characteristics like intertemporal discount rates, cognitive and non-cognitive skills, adverse childhood experiences, frailty, and parental investments. Even though regional or local factors like health policies, migration, crime and pollution could have affected both CSL and mortality, their exclusion will not significantly bias OLS estimates of β_1 as long as SOB is a good proxy for these unobservables. In addition the effectiveness of CSL will be allowed to vary by state, division and region to capture some these variables' incidence on schooling and mortality.

The use of CSL as instruments for schooling is not new in the economics literature. In the only previous study addressing causality, Lleras-Muney (2004) used CSL as instruments for schooling to identify a statistically significant effect on ten-year mortality between -0.051 and -0.037. These IV-2SLS estimates seem too large relative to the average sample mortality of 0.11. Mortality rates are based on synthetic-cohorts constructed from Census data from 1960 to 1980. Section 5 and Appendix Section A1 further discuss the results from Lleras-Muney (2004) and highlight the main differences with the methodology and findings of this paper.

In a somewhat related topic, Acemoglu and Angrist (2000) use CSL and quarter of birth as instruments for education to estimate private (Mincerian) returns to schooling in the order of 7% and social returns (*i.e.* the income return of the state average education) between 1% and 3%. The study suggests that CSL changes between 1910 and 1940 affected schooling mostly around middle school and high school levels, with little or no effect on college education. The authors use this result to support the validity of their instruments, given that unobservable factors like intertemporal discount rates, cognitive and non-cognitive skills, adverse childhood experiences, frailty, and parental investments should also be strongly correlated with college attendance. In addition, this result has important implications on how to interpret the IV estimates of this section.

²⁵ Of the 242,276 observations and 38,348 deaths in the restricted sample, 86,295 observations and 29,792 deaths occur between the ages of 51 and 82. For adults aged 25 to 50, only the correlation between education and mortality can be assessed as there appear to be no valid instruments available.

Finally, Angrist and Krueger (1991) estimate Mincerian returns to schooling using quarter of birth and compulsory school attendance laws as instruments. Their IV estimates are not statistically significantly different from those based on OLS.

Following Lleras-Muney (2004) and Acemoglu and Angrist (2000), education is instrumented with the years of attendance mandated by schooling laws. In every state CSL determine the maximum age by which children must enter school and the minimum age at which they may leave, usually the earliest age they can obtain a work permit. Following these laws from 1915 to 1939, the variable of compulsory years of schooling (CYS) is defined as the difference between the minimum exit age and the maximum entry age. Like previous studies, changes in CSL that were enforced only for one year are ignored. Unlike previous studies, individuals are not matched to mandatory entry and exit ages in place when they were 14 years old, as this ignores the changes in CSL that took place before age 14 and affected schooling decisions. Nevertheless, Acemoglu and Angrist (2000) and Lleras-Muney (2004) justify it on the basis that 14 is the lowest common dropout age across states. An alternative and more appealing specification takes into account the entire history of CSL throughout the 1915-1939 period. Thus, the years of schooling that individuals must obtain before they can join the labor force are contingent on the laws prevalent in their state of birth at the time their decisions to start and leave school took place.

To illustrate how CSL prevalent between 1915 and 1939 translate into CYS for cohorts born between 1901 and 1932, Figure 7 provides the example of the state of California. In 1918, the maximum entry age to school increased one year to 8 and the minimum exit age decreased one year to 14. The first cohort experiencing the exit age reduction, born in 1904, had the possibility to drop out after turning 14 in 1918. Older cohorts had to wait until age 15 instead, so CYS falls from 8 to 7 years for people born in 1904. Since the leaving age was maintained at 14 only until 1920 and then raised back to 15, only individuals born between 1904 and 1907 experienced a one-year decrease in CYS. Similarly, the first cohort experiencing the increase in entry age was born in 1911. After turning 7 years of age in 1918, these individuals had the option to enroll until age 8. Older cohorts, on the contrary, were forced to enroll at age 7. Since no other changes in CSL took place, CYS falls from 8 to 7 for individuals born in 1911 or after. By contrast, Acemoglu and Angrist (2000) and Lleras-Muney (2004) impute a cohort born in 1904 with 6 CYS because entry and exit ages in 1918 were 8 and 16, respectively. This specification erroneously assumes that the 1918 increase in entry age affected the 1904 cohort.

[Figure 7]

The IV-2SLS model to estimate the impact of education on mortality is specified as follows:

$$E_i = \gamma_0 + \gamma_1'Z_i + \gamma_2'X_i + v_i \quad (1^{\text{st}} \text{ IV Stage})$$

$$D_i = \alpha_0 + \alpha_1\hat{E}_i + \alpha_2'X_i + \varepsilon_i \quad (2^{\text{nd}} \text{ IV Stage})$$

The first IV stage regresses years of completed education for individual i (E_i) on a vector of instruments (Z_i) and a set of controls (X_i) comprised of the natural logarithm of per capita income (*Log income*), and dummies for: female gender, state of residence, 1902-1932 cohorts (1901 cohort is the reference category), cohort and female interactions, race (black, Hispanic, and other race; white is the reference category), marital status (single, widowed and divorced; married is the reference category), urban residence, employment status (unemployed, disabled, out of the labor force; employed is the reference category), and inter-state migration. In both stages data is weighted by the individual sampling weight. Estimation was carried out by GMM using White's robust standard errors (reported in parentheses) to correct for heteroskedasticity.²⁶

Table 5 presents the IV-2SLS estimates of α_1 using four different sets of instruments: i) compulsory years of schooling dummies (*CYS*); ii) *CYS* interacted with region of birth dummies (North, South, East or West; *CYS * Region*); iii) *CYS* interacted with division of birth dummies (9 US Census Bureau divisions; *CYS * Division*); and iv) *CYS* interacted with SOB (*CYS * SOB*). All specifications include a continuation school dummy (*Continuation_i*) indicating if working children were required to study on a part-time basis. The rationale for including interacted variables instead of *CYS* dummies is to allow the effectiveness of CSL to vary by region, division or state. Relaxing the assumption of a constant effectiveness of CSL throughout the country also mitigates the problem of omitted variable bias from regional or local factors like health policies, migration, crime and pollution. Moreover, SOB dummies can be excluded from the 2nd IV stage since estimates from Table 4 are not significantly affected by their inclusion. This suggests that any residual effect of SOB on mortality should mainly operate through its impact on education once state of residence and inter-state migration have been controlled for. The potential drawback of including interacted

²⁶ Since Breusch-Pagan, White, Koencker-Basset and Pagan-Hall tests detected the presence of heteroskedasticity, the IV GMM estimator is more efficient than the standard IV-2SLS estimator. Baum, Schaffer and Stillman (2003) describe the properties of each estimator in detail.

instruments, as highlighted by Staiger and Stock (1997), is that IV-2SLS estimates would be biased if such instruments are not valid. Table 5 presents several tests to address this concern.

The first property a valid instrument must satisfy is a significant correlation with the potentially endogenous variable. Two tests evaluate the null hypothesis of a null correlation between schooling and the set of instruments. The first is the F-statistic on the joint significance of instruments in the 1st stage and secondly, Stock and Yogo's (2001) test for weak instruments based on Caigg-Donald's statistic. In all cases, the F-statistic has a p-value below 1% and is greater than 10. This threshold was suggested by Staiger and Stock (1997) as a rule of thumb to detect weak instruments. Stock and Yogo (2001) propose a more formal test which, instead of relying on this rule of thumb, compares Caigg-Donald's statistic to a critical value which depends on how a weak instrument is defined. Under the reported critical values, weak instruments would bias IV estimates by more than 5% relative to OLS estimates.²⁷ This test suggests that *CYS * SOB* is a set of weak instruments when IV regressions are carried out separately by gender and *CYS* is a set of weak instruments for females.

The second property of a valid instrument is its lack of correlation with the 2nd stage error term (ϵ). To test this null hypothesis, Table 5 presents an overidentifying restrictions test in the form of Hansen's J-statistic which consists of the value of the GMM objective function minimized by IV estimates. Under the null, the moment conditions imposing exogeneity between the set of instruments and the 2nd stage residual should be sufficiently close to zero. A rejection of this null would suggest the instruments are not truly exogenous or they have been incorrectly excluded from the 2nd stage regression. All female IV regressions or any IV regression based on *CYS * Region* fail to reject this hypothesis. As a result, *CYS * Region* can be deemed the most appropriate choice of instruments to carry out further IV regressions.

Lastly, Hausman's indirect test of endogeneity is presented in Table 5 to assess the hypothesis of a null correlation between schooling and the 2nd stage error term (ϵ).²⁸ All male IV regressions or any IV regression based on *CYS* or *CYS * Division* fail to reject this hypothesis. However, note that this is not a direct test of endogeneity since ϵ is unobserved; failure to reject the null only implies that OLS and IV estimates are "sufficiently close" in a statistical sense.

[Table 5]

²⁷ See Stock and Yogo (2001) for a more detailed explanation of this test and of the various types of critical values.

²⁸ The C-statistic (of difference-in-Hansen) version of this test was employed to deal with the problem of heteroskedasticity. See Baum, Schaffer and Stillman (2003).

IV-2SLS estimates based only on CYS are not statistically significant but also not statistically different from those based on regional, divisional and state interactions. Moreover, estimates based on the latter specifications are not statistically significantly different from each other. Given that *CYS * Region* uses 5 instrumental variables but *CYS * SOB* uses 49, over-fitting the IV model leads to a bias towards OLS estimates (as noted by Staiger and Stock (1997)). Under the former specification, an additional year of schooling causes on average a 1.6 percentage point reduction on 11-year mortality from an average of 0.33. This marginal effect is roughly thrice as large as the estimated OLS effect. Since omitted variables positively correlated with education and negatively correlated with mortality (like the discount rate, cognitive and non-cognitive skills, frailty and parental inputs) introduce a downward bias in OLS estimates, but classical measurement error results in an upward (attenuation) bias, this result suggests that the latter problem is empirically more important. An alternative explanation relies on the fact that, as suggested by Acemoglu and Angrist (2000), CSL had their greatest impact around middle school and high school levels. In the NLMS sample, CSL changes affected individuals with 10.9 average years of education; and only 11% of them finished college. As shown in Figure 5, the completion of high school is associated with considerable life expectancy improvements. IV estimates thus identify the effect of schooling on mortality around the completion high school. This local average treatment effect (LATE) is greater than the average marginal effect estimated with OLS and Cox PH models.

Results from the first IV stage are presented in Table 6 disaggregating by gender and race. These first-stage regressions use *CYS * Region* and *Continuation* as instrumental variables and include the usual tests for weak instruments, overidentification and endogeneity. Also, data is weighted by individual sampling weights and White's robust standard errors are reported in parentheses. Continuation laws significantly increased the schooling attainment of blacks.

Requiring working children to study on a part-time basis increased black males' schooling by 0.86 years and black females' schooling by 0.57 years. Moreover, the effect of CYS on education varies across regions. CSL in the South seem to have been ineffective as small, statistically insignificant and even negative effects were found across demographic groups. Since no estimate is statistically significant for black females, poor schooling quality and segregation could have undermined the effectiveness of CSL to raise black girls' attainment. For black males, they were only effective in the Northern and Eastern regions where an additional CYS raised schooling by 0.27

years.²⁹ However, the lack of statistical significance in the West is due to a very small sample size. Moreover, the high marginal effects for black males suggest that the laws were most binding for them as they would unlikely obtain a schooling level beyond CYS absent any CSL change. For whites, raising compulsory education by one year increases average schooling by 0.13 and 0.11 years in the North and East but this effect increases to 0.21 in the West. The tests suggest the presence of weak instruments only in the case of black females. The overidentification tests fail to reject the presence of exogenous instruments and endogeneity tests suggest significant differences between OLS and IV estimates only for whites.

[Table 6]

In the second IV stage, the eleven-year death indicator dummy (D_i) is regressed on the predicted value of education (\hat{E}_i) and the set of controls X_i used in the first IV stage. Given the advantages of the original Cox PH model over OLS, a more suitable model to deal with survival data relative to IV-2SLS is the IV Cox PH model:

$$\lambda_i(t = 11; \hat{E}_i, X_i) = \theta_t e^{\kappa_1 \hat{E}_i + \kappa_2' X_i} \quad (\text{IV Cox PH})$$

Variable $\lambda_i(t; \hat{E}_i, X_i)$ captures the t -year hazard rate as in (Cox PH), but the potentially endogenous schooling variable E_i is replaced by its predicted value from the 1st stage \hat{E}_i .

Table 7 presents the estimates of the effect of education on eleven-year mortality for people aged between 51 and 82 years based on the OLS, Cox PH, IV-2SLS and IV Cox PH models. To facilitate comparisons across models, the transformed Cox PH level effects $\psi_1 \lambda$ and $\kappa_1 \lambda$ are reported. Overall, there exists a negative and statistically significant effect of schooling on eleven-year mortality in the order of -0.005 (OLS and Cox PH), -0.016 (IV-2SLS) and -0.021 (IV Cox PH). With an average mortality of 0.34 and an average schooling of 10.85 years, these effects imply an elasticity of mortality with respect to education of -0.5 (IV Cox PH), -0.4 (IV-2SLS) and -0.2 (OLS and Cox PH).³⁰

[Table 7]

²⁹ For reference, the average education level of blacks and whites is 8.3 and 11.3 years respectively. If education is instrumented with CYS alone, the coefficient is 0.12 with a standard error of 0.014. Lleras-Munery's (2004) corresponding estimate is 0.05 with a standard error of 0.007.

³⁰ The mortality elasticity with respect to education ξ is computed with the formula: $\xi \equiv \frac{\partial D}{\partial E} \frac{\bar{E}}{\bar{D}}$, where $\frac{\partial D}{\partial E}$ is the marginal effect estimated in each model and (\bar{E}, \bar{D}) denote average schooling and mortality, respectively.

There are three non-mutually exclusive reasons why IV-2SLS and IV Cox PH estimates are at least twice as large as OLS and Cox PH estimates. First, classical measurement error could be empirically more important than the omitted variables problem.³¹ Second, IV estimates could identify the local average treatment effects (LATE) around high school. Third, non-classical measurement could affect estimates in all regression models. Appendix Section A2 shows that the presence of non-classical measurement errors in NLMS data from unobserved quality of schooling and selective attrition of institutionalized individuals results, under two assumptions, in OLS and IV-2SLS estimates being the lower and upper bounds, respectively and in absolute value, of the true population parameter β_1 . These assumptions are: i) the bias from omitted variables does not dominate the non-classical measurement error biases of OLS estimates and ii) the quality of schooling bias dominates the selective attrition bias of IV-2SLS estimates. Testing this conjecture with evidence from Table 7 shows that for all gender and race groups except black males NLMS estimates satisfy the inequality: $\hat{\alpha}_1^{IV} < \hat{\alpha}_1^{OLS} < 0$. Black males could potentially have a mixture of a strong upward bias from selective attrition (probably caused by a strong negative effect of CSL on the likelihood of incarceration) and a close to null effect of CSL on education in light of the fact that their IV-2SLS estimate is positive.

The use of Cox PH and IV Cox PH models improves the precision of estimates relative to OLS and IV-2SLS. Duration models, unlike OLS and IV-2SLS models, identify statistically significant effects for blacks. This result is important because, in addition to the measurement errors previously described, estimates for blacks are based on a small sample. All models reveal larger estimates for whites than blacks. Although female estimates appear of smaller magnitude relative to males under OLS and Cox PH models, the opposite is true when estimates are based on instrumental variables. Under the Cox PH and IV Cox PH frameworks, these black-white and male-female marginal effect gaps become statistically significant.

Finally, there is a negative effect of income on mortality across most gender and race specifications. Overall, the reduction in 11-year mortality from doubling per capita household income is 0.032 (OLS), 0.044 (Cox PH), 0.015 (IV-2SLS) and 0.018 (IV Cox PH). In all cases, the resulting income elasticity is at most -0.1, implying that education has a proportionally greater impact on mortality than income. Regardless of race, estimates for males have a larger magnitude relative to females.

³¹ Jenkins (2005) presents the direction of bias from omitted variables and measurement errors in the context of the Cox PH duration model.

4. Longevity Gains of Education

4.1 Life Expectancy and Monetized Longevity Gains

To compute the discounted longevity gains of schooling $V_T(a; E)$ from equation (9), the marginal effect of education on survivorship $\left(\frac{\partial \tilde{S}(t, a; E)}{\partial E}\right)$ is approximated by:

$$\frac{\partial \tilde{S}(t, a; E)}{\partial E} \cong \tilde{S}(t, a; E + 1) - \tilde{S}(t, a; E) \quad (16)$$

The right hand side of equation (16) measures the increase in the probability of surviving to age t from age a caused by an additional year of schooling at level E . Combining equations (9) and (16), the discounted longevity gain of an additional year of schooling can be estimated with:

$$V_L(a; E) \cong \int_a^\infty \{\tilde{S}(t, a; E + 1) - \tilde{S}(t, a; E)\} v(t) e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (17)$$

Since $\tilde{S}(t, a; E)$ can be expressed only in terms of the annual hazard rate $\lambda_1(\tau; E)$ as $\tilde{S}(t, a; E) \equiv \prod_{\tau=a}^t (1 - \lambda_1(\tau; E))$, equation (16) can be computed with estimates on the marginal effect of schooling on 11-year mortality from Section 3. In this regard, Cox PH estimates (ψ_1 and κ_1) have two advantages over OLS (β_1) and IV-2SLS (α_1) estimates. First, they don't need to be scaled down to capture the marginal effect on schooling on annual year mortality because, by construction, they identify a percentage marginal effect which is independent of survival time.³² Second and more importantly, they are less precisely measured and generally not statistically different to Cox PH estimates. Thus, the estimated percentage marginal effects can be used to relate $\lambda_1(\tau; E)$ and $\lambda_1(\tau; E + 1)$ as follows:

$$\lambda_1(\tau; E + 1) = \lambda_1(\tau; E)(1 + \psi_1(\tau; E)) \quad (18)$$

Notice that in principle this Cox marginal effect ψ_1 may vary by age τ and education level E . Also, IV Cox PH estimates can be used instead by replacing $\psi_1(\tau; E)$ with $\kappa_1(\tau; E)$. Hence, the marginal effect of education on survivorship can be estimated with:

³² In other words, Cox percentage marginal effects can be converted into to level marginal effects of schooling on T -year mortality by scaling ψ_1 and κ_1 with the T -year hazard λ_T .

$$\frac{\partial \tilde{S}(t, a; E)}{\partial E} \cong \tilde{S}(t, a; E + 1) - \tilde{S}(t, a; E) = \prod_{\tau=a}^t \lambda_1(\tau; E)(1 + \psi_1(\tau; E)) - \prod_{\tau=a}^t \lambda_1(\tau; E) \quad (19)$$

If the Cox percentage marginal effect is assumed constant throughout the lifecycle ($\psi_1(\tau; E) = \psi_1(E) \forall \tau$) equation (19) simplifies to:

$$\frac{\partial \tilde{S}(t, a; E)}{\partial E} \cong \psi_1(E) \prod_{\tau=a}^t \lambda_1(\tau; E) \quad (19)'$$

Estimates of the survivorship gains from an additional year of schooling by gender and race are shown in the top panels of Figure 8. The left panels are based on Cox PH marginal effects (ψ_1) from Table 4 estimated over the entire sample. These marginal effects in percentage (level) terms are: -2.2% (-0.004) for white males, -2.0% (-0.002) for white females, -0.8% (-0.001) for black males and -0.6% (-0.001) for black females. On the other hand, the right panels are based on IV Cox PH marginal effects (κ_1) from Table 7 estimated over the sample aged 51-82. The corresponding percentage (level) marginal effects are: -12.9% (-0.037) for black females, -9.5% (-0.027) for white females, -4.5% (-0.019) for white males and 1.5% (0.006) for black males. Given the restricted sample over which they were estimated, IV Cox PH marginal effects only last between ages 50 and 80 of the lifecycle. The assumption of constant Cox PH marginal effects over the lifecycle is a conservative one since alternative specifications allowing it to vary across 5-year age groups (as in Figure 6) or 20-year age groups considerably increased longevity gains by at least 80%.³³

[Figure 8]

Discounting survivorship gains by the factor $S(t; E) \equiv e^{-r(t-a)} \tilde{S}(t, a = 25; E)$ and weighting them by $v(t)$ yields the present value of monetized of longevity gains from an additional year of schooling: $v(t) \frac{\partial \tilde{S}(t, a; E)}{\partial E} S(t; E)$. The bottom panels of Figure 8 present the longevity gains profile over the lifecycle by gender and race based on Cox PH (ψ_1) and IV Cox PH (κ_1) marginal effects. Even though survivorship gains are very modest at young ages, monetized gains are not negligible since they are not heavily discounted and the value of a life year is relatively high. Regardless of race and gender, survivorship gains are concave over the lifecycle and peak around age 80. However, the concave profile of the value of a life-year $v(t)$ and discounting cause longevity gains to peak up to

³³ This will be shown in Table 8. For the moment, all marginal effects are also assumed constant across schooling levels but future research will relax this restriction.

20 years earlier. The fact that gains from reductions in mortality peak at the onset age of fatal diseases is consistent with Murphy and Topel (2006).

The area under the discounted longevity gain curves yields the expected present value of longevity gains from an additional year of schooling. Table 8 presents these benefits discounted by 3%, 5%, 7% and 10% interest rates.³⁴ In addition, the table reports the life expectancy gain at age a for an individual with E years of schooling; which is computed as:

$$\Delta LE(a; E) = LE(a; E + 1) - LE(a; E) = \int_a^\infty \left\{ \frac{\partial \tilde{s}(t, a; E)}{\partial E} \right\} dt \quad (20)$$

Longevity and life expectancy gains are based on three types of Cox PH estimates. The first set assumes a constant marginal effect of schooling on mortality throughout the lifecycle. The second set allows this marginal effect to vary across three age groups (25-44; 45-64; 65 or older). The third set consists of IV Cox PH estimates, but since they are based on a sample aged 51 to 82, these marginal effects only last between the ages of 50 and 80. Lastly, those gains based on the association between schooling and mortality without the aid of a regression model to control for other covariates is displayed for comparison. Longevity and life expectancy gains correspond to the median values of $V_L(a; E)$ and $\Delta LE(a; E)$ across schooling levels. These may be seen as the median “reduced-form gains” which include the direct effect from education plus the indirect effect caused by covariates or unobservables through schooling.

[Table 8]

Longevity and life expectancy gains based on a constant marginal effect throughout the lifecycle are less than half the median “reduced form” gains. Hence, there is a substantial indirect effect of schooling working through age, marital status, employment status, urban residence and state of residence covariates. Allowing Cox PH marginal effects to vary across 20-year age groups considerably increases longevity gains by at least 80% but barely affects gains in life expectancy. This suggests that even very modest level reductions in mortality at early ages are highly valuable because, as Figure 6 shows, the percentage mortality reductions are large. Assuming a constant percentage marginal effect throughout the lifecycle ignores this feature because only 2% of total deaths occur before age 45, the period when longevity gains are greatest in percentage terms. Gains based on IV

³⁴ Notice that these discount rates ignore any future gains that may accrue from an increase in real wages over time. If the growth rate of real wages (g) is taken into account, the appropriate discount rate would be $r^* \equiv r - g$.

Cox PH marginal effects are the greatest even though they accrue until age 50. As previously conjectured, IV estimates identify the marginal effect around the completion of high school which is associated with great longevity improvements. The most conservative estimates based on Cox PH marginal effects and a 3% discount rate reveal longevity (life expectancy) gains of: \$5,850 (0.22 years) for white males, \$3,650 (0.16 years) for white females, \$1,850 (0.09 years) for black males, and \$900 (0.06 years) for black females. Since the IV-2SLS estimate of black males is positive (albeit not statistically significant), life expectancy and monetized longevity gains are negative. As the discount rate increases, longevity gains based on IV Cox PH estimates are cut more heavily because they accrue 25 years later than those based on Cox PH estimates.

The analysis of life expectancy gains yields interesting predictions regarding the mortality and schooling disparities across racial groups. Under age-constant Cox PH marginal effects, closing the 1.9-year schooling gap between white and black males 25 years or older would reduce the 2.5-year life expectancy gap by approximately 7% (or 0.2 years). If instead black males enjoyed the higher marginal effects of white males, the gap would shrink 17% by 0.4 years. For females, only 3% of the 2-year black-white life-expectancy gap would shrink if the 1.2-year schooling gap were closed. If black females enjoyed the higher marginal effects of white females instead, the life expectancy gap would shrink 10%.

4.2 Income Gains

To estimate income gains $V_Y(a; E)$ of schooling from equation (10), the marginal contribution of schooling to earnings $(\frac{\partial y(t; E)}{\partial E})$ is expressed in terms of Mincerian returns $R(E)$. In the case when these returns are constant by schooling level ($R(E) = R \equiv \delta_1$), this return can be estimated with the Mincerian regression:

$$\ln(y_i) = \delta_0 + \delta_1 E_i + \delta_2' X_i + \varepsilon_i \quad (21)$$

One way to relax this assumption is allowing returns to vary across schooling levels and interacting E_i with education level dummies EL_i^l ($EL_i^l = 1$ if individual i is at schooling level l). Equation (22) specifies Mincerian returns $R(E \in l) = \pi^l$ that vary across five education levels: high school not

finished, completed high school, some college, completed college and graduate/post-college education.

$$\ln(y_i) = \pi_0 + \sum_{l=1}^5 \pi^l E_i * EL_i^l + \pi_2' X_i + \omega_i \quad (22)$$

In both Mincerian equations the dependent variable is the natural logarithm of per capita household income and the set of controls X_i include: years of age,³⁵ age squared, and dummies for: state of residence, marital status (single, widowed, and divorced; married is the baseline category), an urban residence, employment status (unemployed, disabled, and out of the labor force; employed is the baseline category), and inter-state migration.

OLS and IV-2SLS Mincerian estimates from equations (21) and (22) based on individuals aged 51-82 are shown in Table 9 disaggregating by gender and race.³⁶ Unlike IV-2SLS mortality regressions, the instruments for education are the level of CYS interacted with SOB and the continuation school dummy. IV estimates of δ_1 are remarkably similar under *CYS * Region* and *CYS * SOB* specifications. However, the use of *CYS * Region* to instrument five endogenous variables introduces a severe weak instruments problem.

[Table 9]

Ignoring variation across schooling levels, the average OLS (and IV-2SLS) Mincerian returns are: 8.4% (12.1%) for white females, 7.0% (9.8%) for white males, 5.4% (10.8%) for black females and 5.2% (5.3%) for black males. IV-2SLS returns are significantly greater than OLS and all Hausman tests reject the hypothesis that schooling is exogenous. Consistent with previous studies,³⁷ these results suggest the attenuation bias from classical measurement error may be empirically more important than the upward bias induced by omitted variables like the intertemporal discount rate, parental inputs, cognitive skills and non-cognitive skills. Moreover, estimates based on equation (20) have comparable magnitudes to those found in the extensive literature on Mincerian returns. Under both OLS and IV specifications, whites and females appear to have greater returns relative to blacks and males, respectively. Given that OLS estimates from equation (22) reveal increasing returns to schooling regardless of gender and race, there is a strong composition effect where blacks have lower returns than whites partly because they are more concentrated around lower schooling levels.

³⁵ In these Mincerian regressions, age is used as a proxy for work experience.

³⁶ All regressions use person-specific sample weights and White's robust standard errors.

³⁷ Ashenfelter and Card (1999) provide a comprehensive literature survey on income returns to schooling.

IV-2SLS Mincerian returns at or below the high school level are the only statistically insignificant estimates. In general, OLS and IV-2SLS estimates from equation (22) are not statistically different from each other.

4.3 Longevity-Adjusted Mincerian Returns

By definition, the Mincerian return R is the partial derivative of log income with respect to education:

$$R(t; E) = R(E) = \frac{\partial \ln(y(t; E))}{\partial E} \cong \frac{\partial y(t; E)}{\partial E} \frac{1}{y(t; E)} \quad (23)$$

Solving for $\frac{\partial y(t; E)}{\partial E}$ reveals that the marginal income gain defined in equation (10) is the return on the discounted flow of income over the lifecycle:

$$V_Y(a; E) \cong R(E) \int_a^\infty y(t; E) e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (24)$$

Thus, the total (longevity and income) gains of an additional year of education can be estimated with the sum of equations (17) and (24):

$$V_T(a; E) \cong \int_a^\infty \{[\tilde{S}(t, a; E + 1) - \tilde{S}(t, a; E)]v(t) + R(E)y(t; E)\} e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (25)$$

Using a 3% discount rate, Figure 9 presents the estimates of $V_T(a; E)$ by gender and race at five education levels: high school dropouts (average schooling of 8 years), high school graduates (12 completed years of schooling), college dropouts (average schooling of 13.8 years), college graduates (16 completed years of schooling), and people with post-college or graduate education (average schooling of 17.7 years). Income gains $V_Y(a; E)$ are based on OLS Mincerian returns from equation (22) for individuals 25 years or older.³⁸ Longevity gains, on the other hand, are based on Cox PH marginal effects that vary across 20-year age groups. Constant marginal effects over the lifecycle are not used to take into account the fact that modest level reductions in mortality at early ages translate to large percentage changes which are highly valuable. Similarly, IV Cox PH estimates are not used

³⁸ Using estimates from Table 9 based on individuals aged 51 to 82 years did not affect results noticeably. To be conservative, OLS Mincerian return estimates were chosen instead of IV-2SLS estimates because they have a smaller magnitude and are more precise, especially when returns vary by education level.

since they also ignore longevity gains early in the lifecycle and generally yield greater longevity gains than age-varying Cox PH estimates.

[Figure 9]

The average total (income plus longevity) gains from an additional year of schooling $V_T(a = 25; E)$ are approximately: \$93,000 for white females (\$86,000 income, \$7,000 longevity), \$86,000 for white males (\$75,000 income, \$11,000 longevity), \$53,000 for black males (\$50,500 income, \$2,500 longevity), and \$50,000 for black females (\$47,000 income, \$3,000 longevity). Longevity gains tend to increase by education level because, even though the effect of education on mortality is the same for all schooling levels, the value of a life-year and the SVL rise with E . On the other hand, income gains increase with schooling because both Mincerian returns and the income level rise with E .

Longevity gains $V_L(E)$ can be expressed as a rate of return $\theta(E)$ by scaling them to the metric of OLS Mincerian returns:

$$\theta(E) = R(E) \frac{V_L(E)}{V_Y(E)} \quad (26)$$

Thus, the total rate of return to schooling $\Theta(E)$ is:

$$\Theta(E) \equiv \theta(E) + R(E) = R(E) \left(\frac{V_T(E)}{V_Y(E)} \right) \quad (27)$$

By construction, the ratio of longevity returns to total returns equals the ratio of longevity gains to total gains: $\frac{\theta(E)}{\Theta(E)} = \frac{V_L(E)}{V_T(E)}$; and similarly for income: $\frac{R(E)}{\Theta(E)} = \frac{V_Y(E)}{V_T(E)}$.

Figure 10 presents total returns by gender, race and education level. The average total rate of return of an additional year of education is: 7.1% for white males (6.2% income, 0.9% longevity), 7.0% for white females (6.5% income, 0.5% longevity), 5.6% for black males (5.3% income, 0.3% longevity), and 4.9% for black females (4.6% income, 0.3% longevity). Mincerian returns are increasing but longevity returns are decreasing in schooling. Thus, the human capital attained at low and intermediate schooling levels yields proportionately greater longevity benefits relative to specialized skills acquired at more advanced levels. Also, people who discount the future more heavily obtain proportionately lower longevity gains from education since most of these are delayed until late adulthood.

[Figure 10]

5. Evolution of Longevity and Income Returns to Education: 1940-1990

The identification of individual deaths between 1983 and 1994 in NLMS data shed light on the effect of education on mortality and its resulting life expectancy and monetized longevity gains. However, eleven years is a small window to analyze the evolution of total returns to education. Absent any survey data that identifies individual deaths over several decades, this section constructs mortality rates for synthetic cohorts using the Integrated Public Use Microdata Series (IPUMS) from 1940 to 1990.³⁹

5.1 IPUMS Data Description

The selected IPUMS sample consists of 1% random draws of the decennial rounds collected by the Census Bureau from 1940 to 2000. Like the NLMS sample, only US natives (neither born nor residing in Alaska or Hawaii) 25 years or older are included. This restriction yields a sample size of 7,258,182 individual observations. Each Census round provides information on years of completed education, wage income,⁴⁰ age, race, gender, marital status, employment status, SOB and state of residence. In 1950 education and income are available only for survey respondents labeled sample-line persons, which typically are the household heads. To deal with this substantial loss of data in 1950, regressions use sample-line weights instead of person-specific weights in all Census rounds. Future time periods are not affected by this change because person-specific weights differ from sample-line weights only in 1940 and 1950.

A common alternative to measure mortality when individual deaths are unobserved is to follow synthetic cohorts over time. Following Lleras-Muney (2004), the ten-year death rate at period t for group (g, c, s) comprised of individuals of gender (g), cohort (c) and state of birth (s) is defined as:

$$D_t(g, c, s) \equiv \frac{N_t(g, c, s) - N_{t+10}(g, c, s)}{N_t(g, c, s)} = \frac{\sum_{i=N_{t+10}(g, c, s)}^{N_t(g, c, s)} 1}{\sum_{i=1}^{N_t(g, c, s)} 1} \quad (28)$$

³⁹ The data are comprised by the 1940, 1950, 1960 and 2000 1% samples, the 1970 1% Form 2 State Sample, and the 1980 and 1990 1% Metro Samples; all available at <http://usa.ipums.org/usa/>.

⁴⁰ All dollar values are converted to 2004 USD.

The variable $N_t(g, c, s)$ denotes the size of group (g, c, s) at the start of any decade t between 1940 and 1990. The numerator $\sum_{i=N_{t+10}(g, c, s)}^{N_t(g, c, s)} 1$ is an estimate of the number of casualties in the group during decade t , and the denominator $N_t(g, c, s)$ measures the group's risk set, or number of survivors starting period t .

Three important features distinguish death rate $D_t(g, c, s)$ from the NLMS death indicator D_i . First, the former rates can be negative, especially those of younger cohorts which have the lowest mortality rates. In the extreme case of a zero death rate for a given cohort, group size does not change over time. If sampling is random, half of the estimated death rates will be negative to yield a zero average rate. In principle, this should not be a serious concern as long as negative and positive death rates average out around the true value. Alternatively, if death rates are truncated (negative values dropped) or censored (negative values set to zero), true mortality rates will be overestimated.

The second feature is more serious, since death rates cannot be computed for groups that are not observed in subsequent Census rounds. With data from 1940 to 2000, ten-year death rates are available only up to 1990. Mortality measures in 2000 are not identified because $N_{2010}(g, c, s)$ is missing for all groups. To get a sense of the magnitude of these problems, the top panel of Figure 11 shows the fraction of death rates in each five-year age group that are negative, positive, and missing from unobserved $N_{t+10}(g, c, s)$.

[Figure 11]

Negative death rates are more common at younger ages. Consistent with the previous conjecture, approximately half of the death rates between 25 and 39 years of age are negative. This fraction drops to 20% by age 50, and falls below 10% after age 60. On the other hand, missing cohorts in later rounds is a problem arising exclusively after age 80. Overall, 72% of death rates are positive, 26% are negative and only 2% have $N_{t+10}(g, c, s)$ missing.

The third and possibly most critical feature of $D_t(g, c, s)$ is proneness to greater non-classical measurement error compared to D_i . A non-classical measurement error may arise from selective attrition if less educated groups are harder to track down by the Census Bureau. The negative correlation between schooling and risk of attrition results in a downward bias of OLS and IV-2SLS estimates that magnifies the effect of education on mortality. Appendix section A2 presents a formal description of this problem.

To improve some of these features, an alternative death rate $\tilde{D}_t(g, c, s)$ is constructed to incorporate the data's sample design:

$$\tilde{D}_t(g, c, s) \equiv \frac{\tilde{N}_t(g, c, s) - \tilde{N}_{t+10}(g, c, s)}{\tilde{N}_t(g, c, s)} = \frac{\sum_{i=N_{t+10}(g, c, s)}^{N_t(g, c, s)} \omega_i^t(g, c, s)}{\sum_{i=1}^{N_t(g, c, s)} \omega_i^t(g, c, s)} \quad (29)$$

The term $\tilde{N}_t(g, c, s) \equiv \sum_{i=1}^{N_t(g, c, s)} \omega_i^t(g, c, s)$ is the sum of individual sampling weights $\omega_i^t(g, c, s)$ over the $N_t(g, c, s)$ members of group (g, c, s) in period t . As the bottom panel of Figure 11 shows, this alternative specification does not affect the second feature but significantly reduces the fraction of negative death rates. With this specification, 85% of death rates are positive, 13% are negative and 2% are missing from unobserved $N_{t+10}(g, c, s)$. An additional advantage of using \tilde{D}_t instead of D_t is a better fit to ten-year mortality rates from the HMD. This point is illustrated in Figure 12, which includes the censored and truncated versions of \tilde{D}_t for comparison. To avoid the problem of age-composition differences over time, death rates are age-standardized by the population shares in 1990.

[Figure 12]

The HMD age-adjusted ten-year mortality rate is steadily decreasing from 1940 to 1990 by approximately one percentage point every decade.⁴¹ Death rate \tilde{D}_t overestimates the HMD rate by six percentage points in 1950 and 1960, underestimates it by six percentage points in 1970 and is remarkably close to it in 1940, 1980 and 1990. Death rate $D_t(g, c, s)$ has a poorer fit relative to $\tilde{D}_t(g, c, s)$, especially in the years 1950, 1980 and 1990. Censored and truncated death rates, on the other hand, overestimate the HMD rate in all years except 1970.

5.2 Longevity and Income Returns to Education: 1940-1990

To analyze the evolution of the longevity returns of education with IPUMS data, consider the following OLS regression models carried out separately by Census round t :

⁴¹ For the years 1933 to 1969, HMD deaths include US residents and non-residents. Only residents are included from 1970 onward.

$$\tilde{D}_t(g, c, s) = \alpha_0 + \alpha_1 E_{it}(g, c, s) + \alpha_2' X_{it}(g, c, s) + \varepsilon_{it}(g, c, s) \quad (30)$$

$$\tilde{D}_t(g, c, s) = \gamma_0 + \gamma_1 \bar{E}_t(g, c, s) + \gamma_2' \bar{X}_t(g, c, s) + \bar{\varepsilon}_t(g, c, s) \quad (31)$$

Dependent variable $\tilde{D}_t(g, c, s)$ is the ten-year death rate defined in equation (29), $E_{it}(g, c, s)$ is the years of schooling for individual i of group (g, c, s) , and $X_{it}(g, c, s)$ is a vector of controls including dummies for: five-year age group,⁴² state of residence, marital status (single, widowed, and divorced; married is the baseline category), employment status (unemployed, disabled, and out of the labor force; employed is the baseline category), and, where applicable, female, female interacted with age group, and race (black and other non-white race; white is the baseline category). Variables $\bar{E}_t(g, c, s)$ and $\bar{X}_t(g, c, s)$ denote the group averages of $E_{it}(g, c, s)$ and $X_{it}(g, c, s)$, respectively. Finally, $\varepsilon_{it}(g, c, s)$ captures the set of unobservable variables affecting $\tilde{D}_t(g, c, s)$ and $\bar{\varepsilon}_t(g, c, s)$ denotes its group average.

If $\tilde{D}_t(g, c, s)$ measures mortality with classical error, equation (31) becomes the aggregated version of the OLS model previously described in Section 3 and $\varepsilon_{it} = u_{it}, \beta_k = \gamma_k, k = 0, 1, 2$.

$$D_{it} = \beta_0 + \beta_1 E_{it} + \beta_2' X_{it} + u_{it} \quad (\text{OLS})$$

This model requires standard errors to be clustered, resulting in less efficient OLS and IV-2SLS estimates of γ_1 compared to those based on individual data (like equation (30)). Still, it has the same asymptotic properties of model (OLS) provided measurement error in $\tilde{D}_t(g, c, s)$ is classical. In light of the previously described features, this assumption is questionable. An alternative approach is to use OLS and IV-2SLS estimates which, although inconsistent, may set a lower bound for the true effect of education on mortality. In addition, if measurement error does not vary greatly over time the marginal effect trends will scarcely be affected.

With this purpose in mind, four estimates are employed: i) the OLS estimate of γ_1 in equation (31) based on grouped data ($\hat{\gamma}_1^{OLS}$; *OLS Grouped*); ii) the OLS estimate of α_1 in equation (30) adjusted for attenuation from aggregation ($\hat{\alpha}_1^{ADJ}$; *OLS Individual (adjusted)*); iii) the IV-2SLS estimate α_1 in equation (30) based on individual data ($\hat{\alpha}_1^{IV}$; *IV-2SLS Individual (unadjusted)*); and iv) the IV-2SLS estimate of α_1 in equation (30) adjusted for attenuation from aggregation ($\hat{\alpha}_1^{IV-ADJ}$;

⁴² In the case of IV-2SLS regressions, age group dummies are replaced by 1901-1932 cohort dummies because CYS are not defined for other cohorts.

IV-2SLS Individual (adjusted)).⁴³ Appendix Section A4 presents the formal definition of these estimates and their asymptotic properties showing they are valid lower bounds. For reporting purposes, IV-2SLS estimates of equation (31) are ignored since they are not statistically significant and their clustered standard errors are too large to yield any interesting conclusions. Moreover, they have the same probability limit as $\hat{\alpha}_1^{IV-ADJ}$. Likewise, OLS estimates from equation (30) not adjusted for aggregation bias are omitted because, although greatly attenuated, their trends are very similar to those of $\hat{\alpha}_1^{ADJ}$. Trends of these four estimates and their 90% confidence intervals are shown in Figure 13 for the period 1940-1990. OLS estimates are based on individuals 25 years or older and IV-2SLS estimates can only be obtained for cohorts born between 1901 and 1932.

[Figure 13]

Both OLS and IV-2SLS estimates present a remarkably similar pattern throughout the fifty-year period despite being based on samples with different age compositions. In absolute value, estimates are falling from 1940 to 1970 and rapidly rising since. It is unclear why estimates become positive in 1970. The reason may be related to the fact that $\tilde{D}_t(g, c, s)$ significantly underestimates the HMD ten-year mortality rate in that period. Grouped OLS estimates tend to be more attenuated relative to $\hat{\alpha}_1^{ADJ}$, although these differences are not statistically significant. Adjusted IV-2SLS estimates have a significantly smaller magnitude than $\hat{\alpha}_1^{IV}$. These upper-bounds are smaller in magnitude to the estimates of Lleras-Muney (2004) but not statistically different.

Figure 14 further disaggregates these trends by gender and race, analyzing $\hat{\alpha}_1^{IV}$ over time. In general, these demographic groups show trends similar to the ones in Error! Reference source not found., including estimates being positive in 1970 but rapidly falling since. Returns for blacks show greater volatility from decade to decade, especially after 1960 when their marginal effects reach the highest peak (in absolute value). Rising longevity returns from 1940 to 1960 could be associated with racial segregation, and part of the dramatic decline thereafter could be associated with the Civil Liberties Act prohibiting segregation in hospitals and schools. Standard errors are greater for blacks given the low sample size of this group, especially in years 1940 and 1950. Since 1970, longevity returns to schooling rise regardless of gender and race. Interestingly, the trend reversal for black

⁴³ To correct for heteroskedasticity, $\hat{\gamma}_1^{OLS}$ is estimated with GLS from equation (31) using grouped sample-line weights $\tilde{N}_t(g, c, s)$ and White's clustered robust standard errors. Estimates $\hat{\alpha}_1^{ADJ}$, $\hat{\alpha}_1^{IV}$ and $\hat{\alpha}_1^{IV-ADJ}$ based on equation (30) use sample-line weights and White's robust standard errors to correct any remaining non-sphericity in the error term.

males in 1980 is consistent with other studies in the literature showing a slowdown in the convergence of income and achievement gaps across race.⁴⁴

[Figure 14]

With the exception of black males, the results in 1980 are fairly consistent with evidence from Table 7 suggesting whites have greater longevity returns than blacks and males enjoy larger returns than females. Even though IPUMS data presents mortality from 1980 to 1990 and NLMS identifies deaths from 1983 to 1994, longevity returns should not be noticeably different across datasets provided they are based on the same age groups and use similar controls. Empirically, this hypothesis cannot be rejected since IPUMS estimates are not statistically different to NLMS estimates. In 1980 the IPUMS OLS estimate $\hat{\alpha}_1^{ADJ}$ for people 25 years or older is -0.004 for whites and -0.002 for blacks. Remarkably, these coefficients are nearly identical to OLS estimates based on NLMS data (also -0.004 for whites and -0.002 for blacks). Comparing NLMS estimates to $\hat{\gamma}_1^{OLS}$, IPUMS coefficients have a larger magnitudes (-0.010 for whites and -0.008 for blacks) but are not significantly different.

Turning to Mincerian returns, Figure 15 presents the disaggregated trends by gender and race over the period 1940-2000. OLS estimates are based on Mincerian equation (11) using individual (non-clustered) data and the natural logarithm of annual wage earnings as the dependent variable ($\ln(y_i)$).⁴⁵ The set of controls X_{it} include: years of age, age squared, and dummies for: SOB, state of residence, marital status, and employment status. Regressions are carried out separately by gender, race and Census round for individuals aged 25 years or more using person-specific sample-line weights and White's robust standard errors.

[Figure 15]

Mincerian return trends are remarkably similar across gender and race groups decreasing in 1940, and rising steadily after 1950. As documented by Katz & Murphy (1992), in 1980 there is an accelerated growth rate of monetary gains from specialization caused by skill-biased technological change. Since 1950, the returns of white females are larger relative to white males but smaller compared to black females. Even though black males show historically the smallest returns in the

⁴⁴ See Neal (2005).

⁴⁵ Wage values equal to zero are dropped.

period, they converge to the level of white males in 1980 and even surpass them in 2000. Since 1950 black females have enjoyed the highest income return level and growth rate in the population. Selection bias does not explain these trends in view of the fact that results are robust to correction for entry into employment using a Heckman two-step model. The selection-corrected results are not significantly different to those found in Figure 15.⁴⁶

OLS Mincerian returns are notably similar across datasets despite being derived from different income measures. Comparing IPUMS OLS estimates in 1980 to NLMS estimates in 1983 from Figure 15 (in parentheses) reveals: 6.2% (7.0% OLS, 9.8% IV-2SLS) for white males, 7.0% (8.4% OLS, 12.1% IV-2SLS) for white females, 5.7% (5.2% OLS, 5.3% IV-2SLS) for black males and 9.0% (5.4% OLS, 10.8% IV-2SLS) for black females. These similarities arise even though the NLMS sample includes individuals aged between 51 and 82 years, while the IPUMS sample incorporates people 25 years or older. Overall, comparisons across datasets provide consistency and robustness to the results.

6. Discussion and Conclusions

This paper exploited exogenous variation in education induced by compulsory schooling laws implemented in the US between 1925 and 1939 to identify a negative and statistically significant effect of education on eleven-year mortality in the order of -0.005 (OLS and Cox PH), -0.016 (IV-2SLS) and -0.021 (IV Cox PH) using data from the NLMS. These effects translate into life expectancy increases at age 25 of 0.28 and 0.40 years, respectively. A lifetime expected utility maximization model developed by Murphy and Topel (2006) served as the theoretical framework to monetize longevity gains. Using a 3% discount rate, age-varying Cox PH percentage marginal effects and a SVL of \$5 million, the longevity (and life expectancy) gains of an additional year of schooling amount to: \$10,700 (0.24 years) for white males, \$6,700 (0.19 years) for white females, \$3,000 (0.09 years) for black females, and \$2,700 (0.10 years) for black males. Scaling these gains to the metric of Mincerian returns yields an average total (income and longevity) rate of return of: 7.1% for white males (6.2% income, 0.9% longevity), 7.0% for white females (6.5% income, 0.5% longevity), 5.6%

⁴⁶ Marital status was used as the exclusion restriction. Although omitted, these results are available upon request.

for black males (5.3% income, 0.3% longevity), and 4.9% for black males (4.6% income, 0.3% longevity). Regardless of gender and race, longevity gains are long-lived and crest around age 50, just before the onset of fatal diseases. As schooling level rises, longevity gains decrease in proportion to income gains.

The reasons why IV-2SLS and IV Cox PH estimates are larger than OLS and Cox PH estimates are threefold. First, classical measurement error could be empirically more important than the omitted variables problem in mortality and income regressions. Second, IV estimates could be identifying the local average treatment effects (LATE) around the completion of high school which induces greater-than-average longevity improvements. Third, all estimates could suffer from non-classical measurement error caused by selective attrition and unobserved quality of education. In this case, NLMS OLS and IV-2SLS estimates could delimit respectively a lower and an upper bound on the true effect of education on mortality under some assumptions. IPUMS estimates, on the other hand, would provide unambiguous upper bounds of this effect through IV-2SLS. Estimates for blacks, and especially black males, appear to be more severely affected by this source of bias.

The analysis of long-run trends in income and longevity returns with IPUMS data reveals marked similarities across race and gender groups. Longevity returns sharply fall in 1960 but rapidly rise after 1970. With the exception of black males, the growing trend continues until 1990. On the other hand, Mincerian returns decrease in 1940, but steadily rise after 1950. The accelerated growth rate of monetary gains from specialization in 1980 caused by skill-biased technological change is prevalent regardless of race and gender. Even though black males show historically the smallest income returns in the period, they converge to the level of white males in 1980 and even surpass them in 2000. Since 1950 black females have enjoyed the highest level and growth rate of Mincerian returns. Mortality and income estimates are remarkably similar across NLMS and IPUMS datasets in spite of different death rate and income measures, providing consistency and robustness to the results.

The identification of a causal impact on longevity indicates that schooling has a human capital content embodied in factors ultimately related to mortality, such as health investments, smoking, eating and drinking habits or attitudes towards risk. Further research on this area will permit a better understanding on how schooling affects the evolution of racial and gender differences in mortality. Of particular importance is the study of the specific channels through which education improves health and ultimately leads to greater longevity. Datasets like the National Health Interview Survey (NHIS) and the Health and Retirement Study (HRS) may be useful for this.

One form of human capital embedded in schooling is the capacity to reveal information about latent cognitive and non-cognitive skills. The signaling hypothesis, which emphasizes this as the leading role of schooling, is refuted with evidence that schooling improves non-market outcomes. Complementarities between education and other forms of human capital like health and longevity indicate people's skills improve with schooling. Instead of signaling information, the prominent role of education becomes its capacity to enhance the ability of processing information. The present study is one example of how people through education learn to invest in other forms of human capital and collect considerable returns. Better understanding on the capacity of education to improve knowledge and skills will help explain why formal schooling is an essential determinant of market and non-market outcomes in information-based economies; and ultimately, why it is a necessary condition for development.

"Death is a fact. All else is inference." - William Farr (1807–1883)

Bibliography

- Acemoglu, D., & Angrist, J. (2000). How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws. *NBER Macroeconomics Annual* , 15, 9-59.
- Angrist, J. D., & Krueger, A. B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *The Quarterly Journal of Economics* , 106 (4), 979-1014.
- Ashenfelter, O., & Card, D. (1999). The Causal Effect of Education on Earnings. In *Handbook of Labor Economics* (Vol. 3). Amsterdam: Elsevier.
- Baum, C. F., Schaffer, M. E., & Stillman, S. (2003, February). Instrumental Variables and GMM: Estimation and Testing. *Working Paper No. 545* .
- Cox, D. (1972). Regression Models and Life Tables (with Discussion). *Journal of the Royal Statistical Society, Series B* , 187-220.
- Cunha, F., & Heckman, J. J. (2007). The Technology of Skill Formation. *American Economic Review* , 97 (2), 31-47.
- Cutler, D., & Lleras-Muney, A. (2006, July). Education and Health: Evaluating Theories and Evidence. *NBER Working Paper No. 12352* .
- Deaton, A., & Paxson, C. (1999, May). Mortality, Education, Income, and Inequality among American Cohorts. *NBER Working Paper No. W7140* .
- Elias, J. (August 2005). *The Effects of Ability and Family Background on Non-Monetary Returns to Education*. The University of Chicago. PhD Thesis, Department of Economics.
- Felitti, V. (2002). The relationship between adverse childhood experiences and adult health: Turning gold into lead. *The Permanente Journal* , 6, 44-47.
- Fuchs, V. (1982). Time Preference and Health: An Exploratory Study. In V. R. Fuchs (Ed.), *Economic Aspects of Health* (pp. 93-120). Chicago: University of Chicago Press.
- Grossman, M. (2003). Education and Nonmarket Outcomes. (E. H. Welch, Ed.) *Handbook of the Economics of Education* .
- Grossman, M. (1975). The Correlation between Health and Schooling. (N. Terleckyj., Ed.) *Household Production and Consumption* .
- Heckman, J., Stixrud, J., & Urzua, S. (2006). The effect of cognitive and Non-cognitive factors in behavioral and labor outcomes. *Journal of Labor Economics* , Volume 24 (Number 3).
- Jenkins, S. P. (2004). *Survival Analysis*. *Unpublished manuscript*. Retrieved from Institute for Social and Economic Research, University of Essex, Colchester, UK:
<http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/pdfs/ec968notesv6.pdf>
- Jenkins, S. (2005). *Survival Analysis*. Essex, United Kingdom: Institute for Social and Economic Research, University of Essex, unpublished Lecture Notes manuscript.

- Kitagawa, E. M., & Hauser, P. M. (1973). *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Cambridge, Mass.: Harvard University Press.
- Lleras-Muney, A. (2005, January). The Relationship between Education and Adult Mortality in the United States. *Review of Economic Studies* .
- Lochner, L., & Moretti, E. (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self Reports. *American Economic Review* (94), 155-189.
- Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.
- Murphy, K. M., & Topel, R. H. (2006). The Value of Health and Longevity. (T. U. Chicago, Ed.) *Journal of Political Economy* , 114, 871-904.
- Neal, D. (2005). Why Has Black-White Skill Convergence Stopped? *NBER Working Papers Series* (no. 11090).
- Peña, P. (2009). The Effects of Exposure to Domestic Violence in Childhood on Human Capital and Earnings in Adulthood. *Working Paper* .
- Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica* , 65 (3), 557-586.
- Stock, J. H., & Yogo, M. (2001, August). Testing for Weak Instruments in Linear IV Regression. *Working Paper* .
- Viscusi, K. W., & Aldy, J. E. (2003). The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World. *Manuscript, Brookings Joint Center Regulatory Studies* .

Datasets

- NLMS. Norman J. Johnson, Charles Lin, William S. Schauman, Laurita Ajagba, and John L. Boies. National Longitudinal Mortality Study: Public Use File Release 3. Suitland, MD: U.S. Census Bureau, June 1, 2008. <http://www.census.gov/nlms/index.html>
- CSL Data 1915-1939. Adriana Lleras-Muney. <http://www.econ.ucla.edu/alleras/research/data.html>.
- IPUMS. Steven Ruggles, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. Integrated Public Use Microdata Series: Version 4.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2008. <http://www.ipums.org/>
- IHIS. Minnesota Population Center and State Health Access Data Assistance Center, *Integrated Health Interview Series: Version 2.0*. Minneapolis: University of Minnesota, 2008. <http://www.ihis.us>

Appendix

A1. Contrasting methodology and results with Lleras-Muney (2004)

The only previous study addressing the causal effect of education on mortality was done by Adriana Lleras-Muney (2004). Her estimated OLS and IV-2SLS effects of education on eleven-year mortality are -0.017 (with standard error of 0.004) and -0.051 (with standard error of 0.026), respectively. Despite their large disparities, these estimates are not statistically different from each other. However, the IV-2SLS estimate seems implausibly large given that average mortality in her sample is only 0.106. Two factors could have generated these unlikely results.

The first is a data problem. Using Census data from the years 1960, 1970 and 1980, Lleras-Muney (2004) cannot identify individual deaths. Instead, data is aggregated at the gender (g), cohort (c) and state of birth (s) level to construct ten-year death rates at period t with the formula:

$$D_t(g, c, s) \equiv \frac{N_t(g, c, s) - N_{t+10}(g, c, s)}{N_t(g, c, s)},$$

where $N_t(g, c, s)$ is the group size in year t . Dependent variable $D_t(g, c, s)$ is a less precise mortality measure relative to a death indicator. If the measurement error in $D_t(g, c, s)$ is classical (uncorrelated with education), estimates should not differ across datasets. However, changes in cohort size $N_t(g, c, s)$ can be correlated with the group's average schooling level. For example, people with less education could be harder to track down by the Census Bureau. This case of selective attrition introduces a non-classical measurement error problem that biases both OLS and IV-2SLS estimates. Since death rates are artificially high, the resulting downward bias causes the true effect of education on mortality to be magnified (overestimated).

Estimates from Section 5 suggest the use of Census-based death rates magnify the effect of schooling on longevity, and evidence from Lleras-Muney (2004) supports this conjecture. When individual mortality data from the NHEFS (National Health and Nutrition Examination Survey I Follow-up Study, 1992) is used, she finds a statistically significant effect of education on ten-year mortality of -0.011 with OLS and a statistically insignificant IV-2SLS estimate of -0.017.⁴⁷ These estimates are closer to those found in the NLMS (Table 7) and remarkably similar to IPUMS estimates (Figure 14). Appendix Section A2 shows that in the presence of selective attrition and

⁴⁷ This sample is comprised of individuals aged 50 to 74 in 1975 and deaths are identified up to 1985.

unobserved schooling quality, IV-2SLS estimates magnify the true effect of education on mortality and can hence be interpreted as upper bounds.

A second problem arises in the coding of CYS. Lleras-Muney (2004) defines the treatment group as the set of cohorts aged 14 in the year when a particular state CSL was modified, where treatment consists of the number of compulsory schooling years prescribed by the new law. Although age 14 is the lowest common dropout age across states, this specification ignores the potential impact of CSL changes on cohorts of younger age. In addition, it falsely imputes CSL changes on entry to people aged 14, even though they only affected younger cohorts. Using this misspecified variable as the instrument for education with clustered NLMS data at the gender, cohort and state of birth level yields a statistically significant estimate of -0.017 (with standard error 0.007), a value greater but not significantly different to the IV-2SLS estimates from Table 6 using individual NLMS data.

A2. Measurement error bias from selective attrition and quality of education

A2.1 NLMS

Suppose the observed mortality measure suffers from selective attrition and schooling is measured with error from unobserved quality:

$$D = D^* + d \quad (\text{A2.1})$$

$$E = E^* + e \quad (\text{A2.2})$$

The death rate D of a particular individual or group at a given time period is characterized by the true mortality rate D^* plus an error term d that represents the risk of attrition. Measurement error d is negative if attrition results in mortality to be unobserved. For example, the exclusion of institutionalized individuals in the NLMS results in the underestimation of mortality for age groups under 60 and over 80 years of age. Likewise, a positive d indicates that selective attrition results in the over-estimation of mortality, like the case of IPUMS death rates based on synthetic cohorts. Analogously, observed education E is equal to the effective years of schooling E^* plus an error term e capturing its unobserved quality. The “nominal” value of schooling E is thus converted to “real” terms (E^*) by adjusting it with quality of education e . Greater values of e imply poorer quality as the gap between “real” E^* and “nominal” E increases. Although quality is the main focus, measurement

error e can be interpreted more broadly as those cognitive, non-cognitive and parental attributes that contribute to the student's learning efficiency. The lack of those skills and inputs would also introduce a gap between E^* and E . Let the true relationship between quality-adjusted education and mortality (controlling for observables X) be given by:

$$D^* = \beta_0 + \beta_1 E^* + \beta_2' X + \varepsilon \quad (\text{A2.3})$$

The true effect of schooling on mortality is captured by $\beta_1 = \frac{\text{Cov}(D^*, E^*)}{\text{Var}(E^*)}$. If D^* and E^* are unobserved but proxied by D and E , the observed regression model presents measurement error from selective attrition and quality of schooling:

$$D = \alpha_0 + \alpha_1 E + \alpha_2' X + u \quad (\text{A2.4})$$

OLS estimates of α_1 based on equation (A2.4) are biased and inconsistent by three separate channels:

$$\text{plim } \hat{\alpha}_1^{OLS} = \frac{\text{Cov}(D, E)}{\text{Var}(E)} = \frac{\text{Cov}(\beta_0 + \beta_1 E^* + \beta_2' X + \varepsilon + d, E)}{\text{Var}(E)} = \beta_1 \frac{\text{Cov}(E^*, E)}{\text{Var}(E)} + \frac{\text{Cov}(\varepsilon, E)}{\text{Var}(E)} + \frac{\text{Cov}(d, E)}{\text{Var}(E)} \quad (\text{A2.5})$$

The first source of bias is measurement error from unobserved quality of schooling:

$$\frac{\text{Cov}(E^*, E)}{\text{Var}(E)} = \frac{\text{Cov}(E^*, E^* + e)}{\text{Var}(E^* + e)} = \frac{\text{Var}(E^*) + \text{Cov}(E^*, e)}{\text{Var}(E^*) + \text{Var}(e) + 2\text{Cov}(E^*, e)} \in (0, 1) \quad (\text{A2.6})$$

This factor leads to an attenuation (*i.e.* upward) bias as long as: $\text{Var}(e) + \text{Cov}(E^*, e) > 0$ and $\text{Cov}(E^*, E) > 0$. This condition states that the negative relationship between measurement error e and effective years of education E^* (more educated people should be less prone to bad quality of schooling) should not be “too large” (greater than the variance of e). Since β_1 is negative, the first source of bias results in: $\beta_1 < \hat{\alpha}_1^{OLS} < 0$.

The second source of bias is endogeneity from omitted variables like cognitive and non-cognitive skills, discount rates, and parental inputs included in the error term ε . To the extent that these unobservables reduce mortality and positively affect education E^* ($\text{Cov}(\varepsilon, E^*) \leq 0$) and schooling quality ($\text{Cov}(\varepsilon, e) \leq 0$), their omission will result in a negative (*i.e.* downward) bias:

$$\frac{Cov(\varepsilon, E)}{Var(E)} = \frac{Cov(\varepsilon, E^* + e)}{Var(E)} = \frac{Cov(\varepsilon, E^*)}{Var(E)} + \frac{Cov(\varepsilon, e)}{Var(E)} \leq 0 \quad (A2.7)$$

The presence of $Cov(\varepsilon, e) \leq 0$ indicates that unobservables ε not only raise effective years of schooling but also improve its quality, thus magnifying the omitted variable bias. In consequence, the second source of bias results in: $\hat{\alpha}_1^{OLS} < \beta_1 < 0$.

The third source of bias is selective attrition, and its direction depends on how mortality is measured. In the case of the NLMS, the exclusion of institutionalized individuals results in an underestimation of death rates for individuals under age 60 or elders older than 75, and consequently, a negative measurement error d . If schooling reduces the risk of institutionalization (moves d closer to zero) then the selective attrition bias will be positive:

$$\frac{Cov(d, E)}{Var(E)} = \frac{Cov(d, E^*)}{Var(E)} + \frac{Cov(d, e)}{Var(E)} \geq 0 \quad (A2.8)$$

This result is analogous to the phenomenon of survivorship bias from unobserved frailty. In consequence, this source of bias results in the inequality: $\beta_1 < \hat{\alpha}_1^{OLS} < 0$.⁴⁸ If the omitted variable bias is not too strong to dominate the two sources of measurement error bias, the OLS estimate will be a lower bound (in absolute value) of the true parameter β_1 .

The standard approach to correct the omitted variable bias problem is IV. Letting variable Z be a valid instrument for education, the probability limit of the IV-2SLS estimate based on equation (A2.4) is:

$$plim \hat{\alpha}_1^{IV} = \frac{Cov(D, Z)}{Cov(E, Z)} = \frac{Cov(\beta_0 + \beta_1 E^* + \beta_2'X + \varepsilon + d, Z)}{Cov(E, Z)} = \beta_1 \frac{Cov(E^*, Z)}{Cov(E, Z)} + \frac{Cov(d, Z)}{Cov(E, Z)} \quad (A2.9)$$

Although $\hat{\alpha}_1^{IV}$ eliminates the second source of bias affecting OLS estimates, the non-classical measurement error sources still persist. The first source of bias, instead of attenuating, magnifies β_1 provided CSL improved the quality of education.

$$\frac{Cov(E^*, Z)}{Cov(E, Z)} = \frac{Cov(E^*, Z)}{Cov(E^* + e, Z)} = \frac{Cov(E^*, Z)}{Cov(E^*, Z) + Cov(e, Z)} \geq 1 \quad (A2.10)$$

⁴⁸ The case where this bias is sufficiently severe to yield positive OLS estimates ($\beta_1 < 0 < \hat{\alpha}_1^{OLS}$) is ignored as it is not supported empirically.

Better quality is associated with a smaller measurement error e . Assuming CSL improved education to some extent ($Cov(E^*, Z) > 0$) and law amendments were complemented with efforts to improve schooling quality ($Cov(e, Z) < 0$) results in a downward bias: $\hat{\alpha}_1^{IV} < \beta_1 < 0$. If CSL changes did not affect schooling quality ($Cov(e, Z) = 0$), this source of bias vanishes as $\frac{Cov(E^*, Z)}{Cov(E, Z)} = 1$.

The second measurement error source is selective attrition which, like the case of OLS, results in a positive (upward) bias if CSL changes decrease the likelihood of incarceration or eventual institutionalization in a health clinic ($Cov(d, Z) \geq 0$).

$$\frac{Cov(d, Z)}{Cov(E, Z)} = \frac{Cov(d, Z)}{Cov(E^*, Z) + Cov(e, Z)} \geq 0 \quad (A2.11)$$

Relative to OLS, this bias is expected to be smaller and weaker under IV-2SLS because the positive effect that Z may have on d (moving d closer to zero) mainly operates through schooling E . If quality of schooling bias dominates selective attrition, the IV-2SLS estimate will be an upper bound of the true effect of education on mortality: $\hat{\alpha}_1^{IV} < \beta_1 < 0$.

In summary, the presence of non-classical measurement errors in NLMS data from unobserved quality of schooling (e) and selective attrition of institutionalized individuals (d) results, under two assumptions, in OLS and IV-2SLS estimates being the lower and upper bounds in absolute value of the true population parameter β_1 . These assumptions are: i) the omitted variable bias does not dominate the non-classical measurement error biases of OLS estimates and ii) quality of schooling bias dominates the selective attrition bias of IV-2SLS estimates.

Assessing this conjecture with evidence from Table 7 shows that for all gender and race groups except black males NLMS estimates satisfy the inequality: $\hat{\alpha}_1^{IV} < \hat{\alpha}_1^{OLS} < 0$.

All races and genders:	$\hat{\alpha}_1^{IV} = -0.016 < \beta_1 < \hat{\alpha}_1^{OLS} = -0.005 < 0$
All males:	$\hat{\alpha}_1^{IV} = -0.014 < \beta_1 < \hat{\alpha}_1^{OLS} = -0.005 < 0$
All females:	$\hat{\alpha}_1^{IV} = -0.017 < \beta_1 < \hat{\alpha}_1^{OLS} = -0.004 < 0$
All whites:	$\hat{\alpha}_1^{IV} = -0.018 < \beta_1 < \hat{\alpha}_1^{OLS} = -0.006 < 0$
All blacks:	$\hat{\alpha}_1^{IV} = -0.006 < \beta_1 < \hat{\alpha}_1^{OLS} = -0.002 < 0$
White males:	$\hat{\alpha}_1^{IV} = -0.014 < \beta_1 < \hat{\alpha}_1^{OLS} = -0.007 < 0$
White females:	$\hat{\alpha}_1^{IV} = -0.023 < \beta_1 < \hat{\alpha}_1^{OLS} = -0.005 < 0$

Black females: $\hat{\alpha}_1^{IV} = -0.021 < \beta_1 < \hat{\alpha}_1^{OLS} = -0.003 < 0$

Black males, on the other hand, could potentially have a very strong upward bias from selective attrition resulting in $\frac{Cov(d,Z)}{Cov(E,Z)} > 0$ and $\beta_1 < 0 < \hat{\alpha}_1^{IV}$. This could be caused by a large effect of CSL in reducing the likelihood of incarceration among blacks. Alternatively, a small selective attrition bias for blacks could be paired with a null effect of CSL on education ($Cov(E^*, Z) = 0$) to yield a positive IV-2SLS estimate for black males. Empirically, a mixture of both causes is suspected in light of the fact that: $\hat{\alpha}_1^{OLS} = -0.001 < 0 < \hat{\alpha}_1^{IV} = 0.008$.

A2.2 IPUMS

When IPUMS data is used, the probability limits of OLS and IV-2SLS estimates specified in equations (A2.5) and (A2.9) remain unaffected. Even though the sources of bias remain the same, the direction of bias from selective attrition is reversed. IPUMS data includes institutionalized individuals but lacks a death indicator, and therefore requires the construction of ten-year death rates using synthetic cohorts. Let mortality be measured with $D_t \equiv \frac{N_t - N_{t+10}}{N_t}$, where N_t is the size of the group in period t . If this group is hard to track down by the Census Bureau decennially, attrition will lead to the underestimation of N_{t+10} and the overestimation of D_t (measurement error d is positive). If the less educated have greater risk of attrition, the artificial increase in mortality rate will be falsely imputed on schooling, resulting in a downward bias ($\frac{Cov(d,E)}{Var(E)} \leq 0$) of OLS estimates. Thus, $\hat{\alpha}_1^{OLS}$ will be a lower bound in absolute value ($\beta_1 < \hat{\alpha}_1^{OLS} < 0$) only the attenuation bias from unobserved schooling quality dominates selective attrition and omitted variable biases. However, the IV-2SLS estimate becomes an unambiguous upper bound of β_1 provided CSL, through their effect on schooling, reduce the risk of attrition: $\frac{Cov(d,Z)}{Cov(E,Z)} \leq 0$.

$$plim \hat{\alpha}_1^{OLS} = \frac{Cov(D,E)}{Var(E)} = \beta_1 \frac{Cov(E^*,E)}{Var(E)} + \frac{Cov(\varepsilon,E)}{Var(E)} + \frac{Cov(d,E)}{Var(E)} \quad (A2.12)$$

$$plim \hat{\alpha}_1^{IV} = \frac{Cov(D,Z)}{Cov(E,Z)} = \beta_1 \frac{Cov(E^*,Z)}{Cov(E,Z)} + \frac{Cov(d,Z)}{Cov(E,Z)} \quad (A2.13)$$

In summary, OLS estimates $\hat{\alpha}_1^{OLS}$ based on IPUMS data, like in the NLMS, fail to be an unambiguous lower bound of β_1 (in absolute value). However and unlike the NLMS, IV-2SLS estimates are still valid upper bounds: $\hat{\alpha}_1^{IV} < \beta_1 < 0$.

A3. Measurement error and aggregation biases with synthetic-cohort mortality measures

To keep the notation simple, suppress the time subscripts and denote group (g, c, s) only with subscript g . If mortality of individual i from group g could be observed, like in the NLMS, the OLS regression models based on individual and clustered data would be:

$$D_{ig} = \beta_0 + \beta_1 E_{ig} + \beta_2' X_{ig} + u_{ig} \quad (\text{OLS})$$

$$\bar{D}_g = \beta_0 + \beta_1 \bar{E}_g + \beta_2' \bar{X}_g + \bar{u}_g \quad (\text{Grouped OLS})$$

Variables $(\bar{D}_g, \bar{E}_g, \bar{X}_g, \bar{u}_g)$ denote the group averages of $(D_{ig}, E_{ig}, X_{ig}, u_{ig})$. If, on the other hand, individual deaths are unobserved and mortality is measured with \tilde{D}_g following synthetic cohorts over time, a measurement error $(\mu_{ig}, \bar{\mu}_g)$ intrinsically arises:

$$\tilde{D}_g = D_{ig} + \mu_{ig} = \bar{D}_g + \bar{\mu}_g \quad (\text{A3.1})$$

In the second equality, \tilde{D}_g is expressed in terms of the group average measurement error $\bar{\mu}_g$. The corresponding regression models available to the econometrician are:

$$\tilde{D}_g = \alpha_0 + \alpha_1 E_{ig} + \alpha_2' X_{ig} + \varepsilon_{ig} \quad (\text{A3.2})$$

$$\tilde{D}_g = \gamma_0 + \gamma_1 \bar{E}_g + \gamma_2' \bar{X}_g + \bar{w}_g \quad (\text{A3.3})$$

The probability limits of OLS estimates of β_1 are:⁴⁹

⁴⁹ Equation (A4.4) can also be expressed as:

$$\text{plim } \hat{\alpha}_1^{OLS} = \frac{\text{Cov}(\tilde{D}_g, E_{ig})}{\text{Var}(E_{ig})} = \frac{\text{Cov}(\beta_0 + \beta_1 E_{ig} + \beta_2' X_{ig} + u_{ig} + \mu_{ig}, E_{ig})}{\text{Var}(E_{ig})} = \beta_1 + \frac{\text{Cov}(u_{ig}, E_{ig})}{\text{Var}(E_{ig})} + \frac{\text{Cov}(\mu_{ig}, E_{ig})}{\text{Var}(E_{ig})}.$$

$$\begin{aligned}
plim \hat{\alpha}_1^{OLS} &= \frac{Cov(\bar{D}_g, E_{ig})}{Var(E_{ig})} = \frac{Cov(\beta_0 + \beta_1 \bar{E}_g + \beta_2 \bar{X}_g + \bar{u}_g + \bar{\mu}_g, E_{ig})}{Var(E_{ig})} = \\
&= \beta_1 \frac{Cov(\bar{E}_g, E_{ig})}{Var(E_{ig})} + \frac{Cov(\bar{u}_g, E_{ig})}{Var(E_{ig})} + \frac{Cov(\bar{\mu}_g, E_{ig})}{Var(E_{ig})} \quad (A3.4)
\end{aligned}$$

$$plim \hat{\gamma}_1^{OLS} = \frac{Cov(\bar{D}_g, \bar{E}_g)}{Var(\bar{E}_g)} = \frac{Cov(\beta_0 + \beta_1 \bar{E}_g + \beta_2 \bar{X}_g + \bar{u}_g + \bar{\mu}_g, \bar{E}_g)}{Var(\bar{E}_g)} = \beta_1 + \frac{Cov(\bar{u}_g, \bar{E}_g)}{Var(\bar{E}_g)} + \frac{Cov(\bar{\mu}_g, \bar{E}_g)}{Var(\bar{E}_g)} \quad (A3.5)$$

Aside from the omitted variables (\bar{u}_g) and selective attrition ($\bar{\mu}_g$) biases affecting $\hat{\gamma}_1^{OLS}$, $\hat{\alpha}_1^{OLS}$ suffers from aggregation bias $\frac{Cov(\bar{E}_g, E_{ig})}{Var(E_{ig})}$. However, $\hat{\alpha}_1^{OLS}$ is more efficient than $\hat{\gamma}_1^{OLS}$ because it is based on individual data and its standard error is not clustered. In addition, this aggregation bias can be corrected by noticing that $\delta_1 = \frac{Cov(\bar{E}_g, E_{ig})}{Var(E_{ig})}$ is the probability limit of $\hat{\delta}_1^{OLS}$ in the regression:

$$\bar{E}_g = \delta_0 + \delta_1 E_{ig} + \delta_2 \bar{X}_{ig} + \xi_{ig} \quad (A3.6)$$

Thus, a more efficient estimate than $\hat{\gamma}_1^{OLS}$ which unlike $\hat{\alpha}_1^{OLS}$ does not suffer from aggregation bias is: $\hat{\alpha}_1^{ADJ} \equiv \frac{\hat{\alpha}_1^{OLS}}{\hat{\delta}_1^{OLS}}$.⁵⁰ By the Continuous Mapping Theorem:

$$plim \hat{\alpha}_1^{ADJ} = \beta_1 + \frac{\frac{Cov(\bar{u}_g, E_{ig})}{Var(E_{ig})}}{\frac{Cov(\bar{E}_g, E_{ig})}{Var(E_{ig})}} + \frac{\frac{Cov(\bar{\mu}_g, E_{ig})}{Var(E_{ig})}}{\frac{Cov(\bar{E}_g, E_{ig})}{Var(E_{ig})}} = \beta_1 + \frac{Cov(\bar{u}_g, E_{ig})}{Cov(\bar{E}_g, E_{ig})} + \frac{Cov(\bar{\mu}_g, E_{ig})}{Cov(\bar{E}_g, E_{ig})} \quad (A3.7)$$

As seen in Appendix Section A2, omitted variables introduce a negative (downward) bias provided $Cov(\bar{u}_g, E_{ig}) < 0$ and $Cov(\bar{u}_g, \bar{E}_g) < 0$. If selective attrition causes the overestimation of mortality and is negatively related to observed schooling ($Cov(\bar{\mu}_g, \bar{E}_g) < 0$ and $Cov(\bar{\mu}_g, E_{ig}) < 0$) then estimates $\hat{\gamma}_1^{OLS}$ and $\hat{\alpha}_1^{ADJ}$ are unambiguously downward biased. The direction of bias of $\hat{\alpha}_1^{OLS}$ cannot be determined because, in addition to these sources of negative bias, it suffers from a positive (attenuation) bias from aggregation, since $\frac{Cov(\bar{E}_g, E_{ig})}{Var(E_{ig})}$ is empirically between zero and one.

⁵⁰ Estimating standard errors with the Delta method, $\hat{\alpha}_1^{ADJ}$ turns out to be more efficient than $\hat{\gamma}_1^{OLS}$. Alternatively calculating standard errors with $SE(\hat{\alpha}_1^{ADJ}) \equiv \frac{SE(\hat{\alpha}_1^{OLS})}{\hat{\delta}_1^{OLS}}$ yields larger estimates than the Delta method but is computationally simpler.

When a valid instrument for education exists (Z_g such that $Z_g \perp (\bar{u}_g, u_{ig})$), IV-2SLS can be used instead of OLS to eliminate the omitted variable bias. The corresponding probability limits of $\hat{\alpha}_1^{IV}$, $\hat{\gamma}_1^{IV}$ and $\hat{\alpha}_1^{IV-ADJ} \equiv \frac{\hat{\alpha}_1^{IV}}{\hat{\delta}_1^{IV}}$ estimates are:

$$plim \hat{\alpha}_1^{IV} = \frac{Cov(\tilde{D}_g, Z_g)}{Cov(E_{ig}, Z_g)} = \frac{Cov(\beta_0 + \beta_1 \bar{E}_g + \beta_2 \bar{X}_g + \bar{u}_g + \bar{\mu}_g, Z_g)}{Cov(E_{ig}, Z_g)} = \beta_1 \frac{Cov(\bar{E}_g, Z_g)}{Cov(E_{ig}, Z_g)} + \frac{Cov(\bar{\mu}_g, Z_g)}{Cov(E_{ig}, Z_g)} \quad (A3.8)$$

$$plim \hat{\gamma}_1^{IV} = \frac{Cov(\tilde{D}_g, Z_g)}{Cov(\bar{E}_g, Z_g)} = \frac{Cov(\beta_0 + \beta_1 \bar{E}_g + \beta_2 \bar{X}_g + \bar{u}_g + \bar{\mu}_g, Z_g)}{Cov(\bar{E}_g, Z_g)} = \beta_1 + \frac{Cov(\bar{\mu}_g, Z_g)}{Cov(\bar{E}_g, Z_g)} \quad (A3.9)$$

$$plim \hat{\delta}_1^{IV} = \frac{Cov(\bar{E}_g, Z_g)}{Cov(E_{ig}, Z_g)} \quad (A3.10)$$

$$plim \hat{\alpha}_1^{IV-ADJ} \equiv plim \frac{\hat{\alpha}_1^{IV}}{\hat{\delta}_1^{IV}} = \beta_1 + \frac{\frac{Cov(\bar{\mu}_g, Z_g)}{Cov(E_{ig}, Z_g)}}{\frac{Cov(\bar{E}_g, Z_g)}{Cov(E_{ig}, Z_g)}} = \beta_1 + \frac{Cov(\bar{\mu}_g, Z_g)}{Cov(\bar{E}_g, Z_g)} = plim \hat{\gamma}_1^{IV} \quad (A3.11)$$

Even though IV-2SLS corrects for omitted variable bias, estimates remain inconsistent as long as measurement error $\bar{\mu}_g$ is correlated with the instrument Z_g . If selective attrition is negatively correlated with education, an increase in CYS should induce more schooling and in turn decrease the likelihood of attrition: $Cov(\bar{\mu}_g, Z_g) < 0$. Therefore, $\hat{\gamma}_1^{IV}$ and $\hat{\alpha}_1^{IV-ADJ}$ will be biased downward and, like $\hat{\gamma}_1^{OLS}$ and $\hat{\alpha}_1^{ADJ}$, bound the true parameter β_1 from below. Since empirically $\hat{\delta}_1^{IV}$ is very close to unity, $\hat{\alpha}_1^{IV}$ is presumably downward biased as well.

In sum, OLS estimates $\hat{\gamma}_1^{OLS}$ and $\hat{\alpha}_1^{ADJ}$ and IV-2SLS estimates $\hat{\gamma}_1^{IV}$, $\hat{\alpha}_1^{IV}$ and $\hat{\alpha}_1^{IV-ADJ}$ are upper bounds (in absolute value) of the true effect of education on mortality: $UB < \beta_1 < 0$, where $UB = \max \{\hat{\gamma}_1^{OLS}, \hat{\alpha}_1^{ADJ}, \hat{\alpha}_1^{IV}, \hat{\alpha}_1^{IV-ADJ}, \hat{\gamma}_1^{IV}\}$. For reporting purposes, IV-2SLS estimates $\hat{\gamma}_1^{IV}$ are ignored since they are not statistically significant and their clustered standard errors are too large to yield any interesting conclusions. Moreover, $\hat{\alpha}_1^{IV-ADJ}$ is a more efficient substitute of $\hat{\gamma}_1^{IV}$ since, as equation (A3.11) shows, they have the same probability limit.

Tables and Figures

Tables

Table 1. Calibration of parameters

		All	White Males	White Females	Black Males	Black Females
Consumer Surplus $\Phi(Z)$		0.64	0.64	0.64	0.64	0.64
Subsistence ratio Z_0/Z		0.10	0.10	0.10	0.10	0.10
IES σ		1.46	1.46	1.46	1.46	1.46
Average $v(t)$		\$174,057	\$190,424	\$173,809	\$125,766	\$105,141
$v(50)$		\$224,630	\$239,618	\$225,558	\$160,998	\$128,490
SVL						
$r = 3\%$		\$ 5,000,000	\$ 5,310,237	\$ 5,109,467	\$ 3,647,945	\$ 3,082,476
$r = 5\%$		\$ 3,562,529	\$ 3,800,285	\$ 3,622,509	\$ 2,688,232	\$ 2,223,562
$r = 7\%$		\$ 2,701,587	\$ 2,888,511	\$ 2,739,165	\$ 2,087,722	\$ 1,703,957
$r = 10\%$		\$ 1,947,759	\$ 2,085,898	\$ 1,970,803	\$ 1,538,699	\$ 1,241,998

The statistical value of life (SVL), the value of a life-year ($v(t)$) and consumer surplus ($\Phi(z)$) are:

$$SVL = \int_a^{\infty} v(t) e^{-r(t-a)} \tilde{S}(t, a; E) dt$$

$$v(t) = w(t; E)[1 + \Phi(z)]$$

$$\Phi(z) \equiv \left[\frac{u(z)}{zu'(z)} - 1 \right] = \frac{1}{\sigma-1} \left[1 - \sigma \left(\frac{z_0}{z} \right)^{1-\sigma^{-1}} \right]$$

Table 2. NLMS sample restrictions

Restriction	N	Excluded	%		Deaths	Excluded	%
Original sample all ages	988,346				89,909		
Age > 24	579,566	408,780	41%		87,242	2,667	3%
US natives, age >24	545,557	34,009	6%		82,654	4,588	5%
US natives, age >24, SOB not missing	242,276	303,281	56%		38,348	44,306	54%
Restricted Sample	242,276	746,070	75%		38,348	51,561	57%

Note: Except for the last row, % denotes the fraction of the sample excluded relative to the previous row total

The exclusion of individuals under 25 years and those born in Hawaii, Alaska or abroad reduces the sample by only 6% (5% of total deaths). Missing state of birth in 56% of the sample is a sizable source of attrition, responsible for the loss of 54% of total identified deaths.

Table 3. NLMS summary statistics with missing and not missing state of birth (SOB) characteristics

Variables	SOB Not Missing			Missing SOB			Mean diff.	%diff.	Mean Test			Std. Dev. Test	
	N	Mean (1)	Std. Dev. (1)	N	Mean (2)	Std. Dev. (2)			T-test	P Value		F-test	P Value
Death	242276	0.16	0.37	303281	0.15	0.35	0.012	8%	12.4	0.00		1.07	0.00
Education	240794	12.0	3.24	302947	12.1	3.40	-0.003	0%	-0.3	0.77		0.91	0.00
Age	242276	46.1	15.88	303281	46.8	16.17	-0.638	-1%	-14.6	0.00		0.96	0.00
Per capita income	226791	18733.0	11584.2	300849	21013.8	12805.5	-2280.7	-11%	-67.6	0.00		0.82	0.00
Average income	226791	31615.1	19114.4	300849	35244.8	21420.6	-3629.8	-10%	-64.8	0.00		0.80	0.00
Female	242276	0.53	0.50	303281	0.54	0.50	-0.004	-1%	-2.9	0.00		1.00	0.78
White	242243	0.87	0.33	301866	0.82	0.38	0.052	6%	53.9	0.00		0.75	0.00
Black	242243	0.09	0.29	301866	0.09	0.28	0.008	9%	10.1	0.00		1.08	0.00
Hispanic	242243	0.02	0.16	301915	0.08	0.27	-0.053	-68%	-91.2	0.00		0.34	0.00
Other Race	242243	0.01	0.09	301866	0.02	0.12	-0.007	-46%	-24.8	0.00		0.54	0.00
Married	240326	0.73	0.45	303076	0.71	0.46	0.021	3%	16.8	0.00		0.96	0.00
Single	240326	0.09	0.29	303076	0.10	0.30	-0.008	-8%	-10.1	0.00		0.93	0.00
Divorced	240326	0.10	0.30	303076	0.11	0.31	-0.007	-7%	-9.1	0.00		0.94	0.00
Widowed	240326	0.08	0.27	303076	0.09	0.28	-0.005	-6%	-6.6	0.00		0.95	0.00
Urban	240507	0.65	0.48	303123	0.68	0.46	-0.038	-6%	29.4	0.00		1.06	0.00
Employed	240794	0.62	0.48	302042	0.61	0.49	0.013	2%	9.6	0.00		0.99	0.00
Unemployed	240794	0.03	0.18	302042	0.04	0.20	-0.008	-19%	-14.9	0.00		0.82	0.00
Disabled	240794	0.02	0.13	302042	0.02	0.13	0.000	2%	1.1	0.25		1.02	0.00
Out of LF	240794	0.33	0.47	302042	0.33	0.47	-0.005	-2%	-4.3	0.00		0.99	0.00
Notes:	Mean P value reports the probability that difference in means (Mean(1)-Mean(2)) is equal to zero.												
	Standard deviation F-test reports the ratio of Std. Dev. (1) / Std. Dev. (2).												
	Standard deviation P value reports the probability that the ratio is equal to one.												

The number of observations (*N*), mean and standard deviation of eleven-year mortality, education, age and other socio-demographic characteristics are reported for two subsamples: 1) state of birth (SOB) missing, and 2) SOB not missing. Mean and standard deviation differences are statistically significant across subsamples but small in magnitude. As main exceptions, non-black racial minorities (*Hispanic* and *other race* dummies) and the unemployed are under-represented in the subsample with SOB not missing.

Table 4. The schooling–mortality gradient: OLS and Cox PH models

Age	ALL			WHITES			BLACKS		
	OLS (β)	COX PH ($\lambda\psi$)	11yr DR (λ)	OLS (β)	COX PH ($\lambda\psi$)	11yr DR (λ)	OLS (β)	COX PH ($\lambda\psi$)	11yr DR (λ)
BOTH GENDERS									
25+	-0.003 (0.0003)***	-0.002 (0.0001)***	0.15	-0.004 (0.0003)***	-0.003 (0.0001)***	0.16	-0.002 (0.0010)	-0.001 (0.0003)***	0.17
51-82	-0.005 (0.0006)***	-0.005 (0.0001)***	0.33	-0.006 (0.0006)***	-0.006 (0.0001)***	0.33	-0.002 (0.0017)	-0.003 (0.0003)***	0.37
25-44	-0.002 (0.0002)***	-0.001 (0.0005)***	0.02	-0.002 (0.0002)***	-0.001 (0.0005)***	0.02	-0.002 (0.0010)**	-0.002 (0.0011)***	0.04
45-64	-0.005 (0.0006)***	-0.004 (0.0002)***	0.15	-0.006 (0.0006)***	-0.005 (0.0002)***	0.14	-0.002 (0.0018)	-0.002 (0.0005)***	0.21
65+	-0.004 (0.0008)***	-0.006 (0.0001)***	0.58	-0.005 (0.0009)***	-0.007 (0.0002)***	0.59	-0.001 (0.0026)	-0.003 (0.0004)***	0.57
MALES									
25+	-0.004 (0.0004)***	-0.003 (0.0001)***	0.17	-0.005 (0.0004)***	-0.004 (0.0002)***	0.18	-0.002 (0.0015)	-0.001 (0.0004)***	0.20
51-82	-0.006 (0.0008)***	-0.002 (0.0002)***	0.39	-0.007 (0.0009)***	-0.009 (0.0002)***	0.39	-0.002 (0.0026)	-0.002 (0.0005)***	0.45
25-44	-0.002 (0.0004)***	-0.002 (0.0006)***	0.03	-0.002 (0.0004)***	-0.002 (0.0006)***	0.03	-0.004 (0.0018)**	-0.003 (0.0013)***	0.05
45-64	-0.005 (0.0008)***	-0.005 (0.0002)***	0.19	-0.006 (0.0009)***	-0.006 (0.0003)***	0.19	0.000 (0.0027)	0.000 (0.0006)**	0.28
65+	-0.004 (0.0012)***	-0.008 (0.0002)***	0.68	-0.005 (0.0013)***	-0.009 (0.0002)***	0.68	-0.002 (0.0040)	-0.002 (0.0006)***	0.67
FEMALES									
25+	-0.003 (0.0004)***	-0.002 (0.0002)***	0.13	-0.003 (0.0004)***	-0.002 (0.0002)***	0.14	-0.002 (0.0013)	-0.001 (0.0004)***	0.14
51-82	-0.004 (0.0008)***	-0.005 (0.0001)***	0.27	-0.005 (0.0009)***	-0.006 (0.0001)***	0.27	-0.002 (0.0024)	-0.003 (0.0003)***	0.31
25-44	-0.001 (0.0003)***	-0.001 (0.0009)***	0.02	-0.001 (0.0003)***	-0.001 (0.0010)***	0.01	-0.001 (0.0010)	-0.001 (0.0020)***	0.03
45-64	-0.004 (0.0008)***	-0.003 (0.0004)***	0.11	-0.004 (0.0008)***	-0.004 (0.0004)***	0.11	-0.004 (0.0025)	-0.003 (0.0008)***	0.16
65+	-0.003 (0.0012)***	-0.004 (0.0002)***	0.52	-0.005 (0.0013)***	-0.006 (0.0002)***	0.52	-0.001 (0.0035)	-0.003 (0.0006)***	0.51

Notes: X+ denotes X years or older

*, ** and *** denote statistical significance at 90%, 95% and 99% confidence levels, respectively

$$D = \beta_0 + \beta_1 E + \beta_2' X + u \quad (\text{OLS})$$

$$\lambda(t = 11; E, X) = \theta_t e^{\psi_1 E + \psi_2' X} \quad (\text{Cox PH})$$

Variable D is the eleven-year death indicator, $\lambda(t; E, X)$ is the t -year hazard rate (with time-dependent hazard factor θ_t), E is years of completed education and X denotes a set of controls including the natural logarithm of per capita income and dummies for: SOB, state of residence, five-year age group, marital status, urban residence, inter-state migration and employment status. Data is weighted by individual sampling weights and OLS estimates report White's robust standard errors.

Table 5. IV-2SLS weak instruments and over-identification tests

	Model Instruments Z	OLS	IV-2SLS CYS	IV-2SLS CYS*Region	IV-2SLS CYS*Division	IV-2SLS CYS*SOB
All	(N=79,146)					
	Education	-0.005 (0.001)***	0.000 (0.007)	-0.016 (0.005)***	-0.013 (0.004)***	-0.012 (0.004)***
Weak instruments tests						
	F-Stat		33.77***	167.13***	92.57***	23.86***
	Craigg-Donald F-Stat		45.60	233.43	127.38	33.03
	Critical value		20.90	18.37	20.74	21.42
Overidentification test						
	Hansen's J-Stat		17.67	4.20	17.72	66.18
	P-value		0.04	0.24	0.02	0.03
Endogeneity test						
	Hausman's J-Stat		0.51	4.91	3.07	4.91
	P-value		0.48	0.03	0.08	0.03
Males	(N=35,570)					
	Education	-0.005 (0.001)***	0.005 (0.008)	-0.014 (0.006)**	-0.013 (0.006)**	-0.008 (0.005)
Weak instruments tests						
	F-Stat		21.87***	101.95***	57.24***	14.71***
	Craigg-Donald F-Stat		30.56	145.12	79.97	20.4
	Critical value		20.90	18.37	20.74	21.42
Overidentification test						
	Hansen's J-Stat		17.9	4.25	15.79	66.70
	P-value		0.04	0.24	0.05	0.03
Endogeneity test						
	Hausman's J-Stat		1.51	1.95	1.99	1.28
	P-value		0.22	0.16	0.16	0.26
Females	(N=43,576)					
	Education	-0.004 (0.001)***	-0.002 (0.012)	-0.017 (0.007)**	-0.013 (0.007)*	-0.015 (0.006)***
Weak instruments tests						
	F-Stat		13.17***	67.76***	37.34***	11.13***
	Craigg-Donald F-Stat		17.01	91.68	50.17	14.6
	Critical value		20.90	18.37	20.74	21.42
Overidentification test						
	Hansen's J-Stat		9.3	1.35	9.02	50.55
	P-value		0.41	0.72	0.34	0.34
Endogeneity test						
	Hausman's J-Stat		0.06	3.00	1.27	3.48
	P-value		0.81	0.08	0.26	0.06

Notes: *, ** and *** denote statistical significance at 90%, 95% and 99% confidence levels, respectively

Data is weighted by the person-specific sample weight

White's robust standard errors are reported in parentheses

The null hypotheses for each test are: Weak instruments tests. $H_0: \text{Cov}(Z, E) = 0$

Overidentification test. $H_0: \text{Cov}(Z, \epsilon) = 0$

Endogeneity test. $H_0: \text{Cov}(E, \epsilon) = 0$

Critical value' refers to Stock & Yogo's (2001) test to reject a bias in IV no greater than 5% relative to OLS

$$D = \alpha_0 + \alpha_1 \hat{E} + \alpha_2' X + \varepsilon \quad (2^{\text{nd}} \text{ IV Stage})$$

The sample is comprised of non-institutionalized US natives (excluding natives or residents from Alaska and Hawaii) aged 51-82. The instruments for education are a continuation school dummy and CYS interacted with SOB dummies. Additional controls (X) include dummies for: state of residence, 1902-1932 cohort (1901 cohort is reference category), marital status (single, widowed and divorced; married is reference category), urban residence, employment status (unemployed, disabled, out of the labor force; employed is the reference category), and inter-state migration. The regression uses person-specific sampling weights and White's robust standard errors (reported in parentheses).

Two tests evaluate the null hypothesis of a null correlation between schooling and the set of instruments. The first is the F-statistic from the 1st stage on the joint significance of instruments and secondly, Stock and Yogo's (2001) test for weak instruments based on Caigg-Donald's statistic. An overidentifying restrictions test evaluates the null hypothesis of lack of correlation between the instrument and the error term in the 2nd stage. This is the Hansen's J-statistic which consists of the estimated value of the GMM objective function minimized by IV estimates. Lastly, Hausman's indirect test of endogeneity assesses the null hypothesis of a null correlation between schooling and the 2nd stage error term (ϵ).

Table 6. The IV 1st Stage: the effect of instruments on schooling

Gender Race Instrument	BOTH			MALES			FEMALES		
	ALL	WHITES	BLACKS	ALL	WHITES	BLACKS	ALL	WHITES	BLACKS
Continuation	0.15 (0.04)***	0.02 (0.04)	0.67 (0.15)***	0.25 (0.06)***	0.12 (0.06)*	0.86 (0.24)***	0.06 (0.05)	-0.08 (0.05)	0.57 (0.18)***
CYS*Region									
CYS North	0.13 (0.01)***	0.13 (0.01)***	0.19 (0.05)***	0.14 (0.02)***	0.13 (0.02)***	0.27 (0.08)***	0.12 (0.02)***	0.13 (0.02)***	0.10 (0.07)
CYS South	0.00 (0.02)	-0.02 (0.02)	0.02 (0.05)	-0.02 (0.02)	-0.04 (0.03)	0.08 (0.08)	0.01 (0.02)	0.01 (0.02)	-0.03 (0.07)
CYS East	0.11 (0.01)***	0.11 (0.02)***	0.17 (0.05)***	0.13 (0.02)***	0.10 (0.02)***	0.28 (0.08)***	0.10 (0.02)***	0.11 (0.02)***	0.07 (0.07)
CYS West	0.21 (0.02)***	0.21 (0.02)***	-0.07 (0.12)	0.22 (0.03)***	0.21 (0.03)***	0.29 (0.18)	0.19 (0.02)***	0.20 (0.02)***	-0.34 (0.15)
N	79146	70225	7049	35570	31730	2972	43576	38495	4076
R-squared	0.27	0.28	0.20	0.27	0.28	0.20	0.24	0.25	0.15
Mean Education	10.85	11.16	8.42	10.87	11.17	7.92	10.83	11.15	8.79
Weak instruments tests (Critical value = 18.37)									
F-Stat	167.1***	146.1***	29.6***	102.0***	81.0***	20.0***	67.8***	66.2***	12.3***
Craigg-Donald F-Stat	233.4	202.0	33.1	145.1	114.3	23.3	91.7	88.8	12.5
Overidentification test									
Hansen's J-Stat	4.20	1.83	3.14	4.25	2.41	3.07	1.35	0.39	2.36
P-value	0.24	0.61	0.53	0.24	0.49	0.55	0.72	0.94	0.67
Endogeneity test									
Hausman's J-Stat	4.91	5.48	0.01	1.95	1.20	0.30	3.00	5.07	1.03
P-value	0.03	0.02	0.92	0.16	0.27	0.58	0.08	0.02	0.31

Notes:

Instruments for education: CYS interacted with region of birth + continuing school dummy

*, ** and *** denote statistical significance at 90%, 95% and 99% confidence levels, respectively

Not SS denotes the number of states whose CYS estimate is not statistically significant at the 90% confidence level

Data is weighted by the person-specific sample weight

White's robust standard errors are reported in parentheses

Critical value' refers to Stock & Yogo's (2001) test to reject a bias in IV no greater than 5% relative to OLS

The null hypotheses for each test are:

Weak instruments tests. $H_0: \text{Cov}(Z, E) = 0$

Overidentification test. $H_0: \text{Cov}(Z, \varepsilon) = 0$

Endogeneity test. $H_0: \text{Cov}(E, \varepsilon) = 0$

$$E = \gamma_0 + \gamma_1'Z + \gamma_2'X + v \quad (1^{\text{st}} \text{ IV Stage})$$

Education (E) is regressed against a set of instruments (Z) and a set of controls (X) that include the natural logarithm of per capita income, and dummies for: state of residence, birth cohort from 1902 to 1932, marital status, urban residence, employment status, inter-state migration and, where applicable, female, cohort interacted with female, and race. The instruments for education are a continuation school dummy and CYS interacted with region of birth. The regression uses person-specific sampling weights and White's robust standard errors (reported in parentheses).

Table 7. Effect of schooling on 11-year mortality: OLS, Cox PH, IV-2SLS and IV Cox PH models

Race		ALL				WHITES				BLACKS			
Model		OLS	COX PH	IV-2SLS	IV COX PH	OLS	COX PH	IV-2SLS	IV COX PH	OLS	COX PH	IV-2SLS	IV COX PH
		(β_1)	($\lambda\psi_1$)	(α_1)	($\lambda\kappa_1$)	(β_1)	($\lambda\psi_1$)	(α_1)	($\lambda\kappa_1$)	(β_1)	($\lambda\psi_1$)	(α_1)	($\lambda\kappa_1$)
BOTH GENDERS													
	Education	-0.005	-0.005	-0.016	-0.021	-0.006	-0.007	-0.018	-0.022	-0.002	-0.002	-0.006	-0.007
		(0.001)***	(0.001)***	(0.005)***	(0.001)***	(0.001)***	(0.001)***	(0.005)***	(0.001)***	(0.002)	(0.000)***	(0.011)	(0.001)***
	Log income	-0.032	-0.044	-0.015	-0.018	-0.033	-0.043	-0.013	-0.017	-0.017	-0.028	-0.012	-0.021
		(0.003)***	(0.001)***	(0.008)*	(0.001)***	(0.003)***	(0.001)***	(0.009)	(0.001)***	(0.009)**	(0.001)***	(0.017)	(0.001)***
	N	79146	79146	79146	79146	70225	70225	70225	70225	7049	7049	7049	7049
	R-squared	0.27		0.27		0.28		0.28		0.20		0.20	
	Mean death	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.40	0.40	0.40	0.40
MALES													
	Education	-0.005	-0.007	-0.014	-0.018	-0.007	-0.009	-0.014	-0.019	-0.001	-0.001	0.001	0.006
		(0.001)***	(0.001)***	(0.006)**	(0.001)***	(0.001)***	(0.001)***	(0.007)**	(0.001)***	(0.003)	(0.000)***	(0.013)	(0.001)***
	Log income	-0.043	-0.060	-0.027	-0.037	-0.041	-0.056	-0.027	-0.037	-0.045	-0.062	-0.049	-0.072
		(0.004)***	(0.001)***	(0.012)**	(0.001)***	(0.005)***	(0.001)***	(0.014)**	(0.001)***	(0.014)***	(0.001)***	(0.023)**	(0.002)***
	N	35570	35570	35570	35570	31730	31730	31730	31730	2973	2973	2973	2973
	R-squared	0.28		0.27		0.29		0.28		0.21		0.20	
	Mean death	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.49	0.49	0.49	0.49
FEMALES													
	Education	-0.004	-0.004	-0.017	-0.023	-0.005	-0.005	-0.023	-0.027	-0.003	-0.003	-0.021	-0.037
		(0.001)***	(0.001)***	(0.007)**	(0.001)***	(0.001)***	(0.001)***	(0.008)***	(0.001)***	(0.002)	(0.000)***	(0.017)	(0.001)***
	Log income	-0.023	-0.030	-0.004	-0.002	-0.026	-0.033	0.000	0.001	0.007	0.004	0.032	0.050
		(0.004)***	(0.001)***	(0.012)	(0.001)***	(0.004)***	(0.001)***	(0.012)	(0.001)	(0.011)	(0.001)***	(0.026)	(0.002)***
	N	43576	43576	43576	43576	38495	38495	38495	38495	4076	4076	4076	4076
	R-squared	0.24		0.24		0.26		0.25		0.17		0.15	
	Mean death	0.29	0.29	0.29	0.29	0.28	0.28	0.28	0.28	0.32	0.32	0.32	0.32
Notes: Instruments for education: CYS interacted with region of birth + continuing school dummy													
*, ** and *** denote statistical significance at 90%, 95% and 99% confidence levels, respectively													
Data is weighted by the person-specific sample weight, and in the case of OLS and IV-2SLS, White's robust standard errors are reported in parentheses													

$$D = \beta_0 + \beta_1 E + \beta_2' X + u \quad (\text{OLS})$$

$$\lambda(t = 11; E, X) = \theta_t e^{\psi_1 E + \psi_2' X} \quad (\text{Cox PH})$$

$$D = \alpha_0 + \alpha_1 \hat{E} + \alpha_2'X + \varepsilon \quad (2^{\text{nd}} \text{ IV Stage})$$

$$\lambda(t = 11; E, X) = \theta_t e^{\kappa_1 \hat{E} + \kappa_2'X} \quad (\text{IV Cox PH})$$

The sample is comprised of non-institutionalized US natives (excluding natives or residents from Alaska and Hawaii) aged 51-82. Variable D is the eleven-year death indicator, $\lambda(t; E, X)$ is the t -year hazard rate (with time-dependent hazard factor θ_t), E is years of completed education and \hat{E} denotes its fitted value from the 1st IV stage. The instruments for education are a continuation school dummy and CYS interacted with region of birth. Additional controls (X) include dummies for: state of residence, 1902-1932 cohort, marital status, urban residence, employment status, and inter-state migration. All regressions use person-specific sampling weights. OLS and IV-2SLS regressions report White's robust standard errors in parentheses.

Table 8. Life expectancy and discounted longevity gains of schooling

		All	Males		Females	
			White	Black	White	Black
Median gains (no regression)*						
Life expectancy gain		0.33	0.44	0.70	0.17	0.44
Longevity gain	$r = 3\%$	\$ 10,884	\$ 13,677	\$ 24,949	\$ 15,914	\$ 5,197
	$r = 5\%$	\$ 5,335	\$ 9,569	\$ 9,448	\$ 7,677	\$ 5,657
	$r = 7\%$	\$ 3,093	\$ 5,585	\$ 3,716	\$ 4,136	\$ 3,082
	$r = 10\%$	\$ 1,213	\$ 2,013	\$ 1,202	\$ 2,342	\$ 1,329
Cox PH Estimate						
Marginal effect (%)		-1.6%	-2.2%	-0.8%	-2.0%	-0.6%
Life expectancy gain		0.15	0.22	0.09	0.16	0.06
Longevity gain	$r = 3\%$	\$ 3,748	\$ 5,857	\$ 1,854	\$ 3,650	\$ 897
	$r = 5\%$	\$ 1,797	\$ 2,883	\$ 1,001	\$ 1,656	\$ 446
	$r = 7\%$	\$ 929	\$ 1,518	\$ 579	\$ 812	\$ 241
	$r = 10\%$	\$ 397	\$ 660	\$ 287	\$ 324	\$ 112
Marginal effect (%) by age group						
25-44		-5.7%	-7.2%	-2.3%	-6.7%	-4.6%
45-64		-3.1%	-3.7%	-0.2%	-4.4%	-1.0%
65+		-1.0%	-1.0%	-0.9%	-1.2%	-0.1%
Life expectancy gain		0.19	0.24	0.10	0.19	0.09
Longevity gain	$r = 3\%$	\$ 7,228	\$ 10,691	\$ 2,730	\$ 6,699	\$ 3,066
	$r = 5\%$	\$ 3,858	\$ 5,834	\$ 1,673	\$ 3,404	\$ 1,832
	$r = 7\%$	\$ 2,201	\$ 3,382	\$ 1,091	\$ 1,852	\$ 1,162
	$r = 10\%$	\$ 1,068	\$ 1,667	\$ 630	\$ 842	\$ 648
IV Cox PH Estimate**						
Marginal effect (%)		-6.0%	-4.5%	1.5%	-9.5%	-12.9%
Life expectancy gain		0.40	0.33	-0.12	0.53	0.86
Longevity gain	$r = 3\%$	\$ 8,382	\$ 7,348	\$ (1,632)	\$ 11,185	\$ 10,846
	$r = 5\%$	\$ 3,517	\$ 3,161	\$ (724)	\$ 4,561	\$ 4,531
	$r = 7\%$	\$ 1,521	\$ 1,396	\$ (330)	\$ 1,920	\$ 1,959
	$r = 10\%$	\$ 456	\$ 429	\$ (106)	\$ 556	\$ 592
Life expectancy at age 25		52.45	49.80	47.33	54.91	52.92
Life expectancy at age 50		27.74	25.14	23.04	30.09	28.35
Average Education (ages 25+)		12.01	12.54	10.61	12.10	10.88
Average Education (ages 50+)		10.85	11.24	7.84	11.13	8.64

Note: * Median discounted longevity gains with respect to years of schooling without a regression controlling for other covariates.

** Marginal effects last between ages 50 and 80 since estimates are based on individuals ages 51-82.

The life expectancy gain of an additional year of schooling is:

$$\Delta LE(a; E) = LE(a; E + 1) - LE(a; E) = \int_a^\infty \left\{ \frac{\partial \tilde{S}(t, a; E)}{\partial E} \right\} dt \quad (20)$$

The present value discounted longevity gains at age a , $V_L(a = 25; E)$, are:

$$V_L(a = 25; E) = \int_a^\infty \left\{ v(t) \frac{\partial \tilde{S}(t, a; E)}{\partial E} \right\} e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (9)$$

$$\frac{\partial \tilde{S}(t, a; E)}{\partial E} \cong \tilde{S}(t, a; E + 1) - \tilde{S}(t, a; E) = \prod_{\tau=a}^t \lambda_1(\tau; E) (1 + \psi_1(\tau; E)) - \prod_{\tau=a}^t \lambda_1(\tau; E) \quad (19)$$

Table 9. Mincerian returns to education by gender and race: OLS & IV-2SLS

Race Education Level	ALL		WHITES		BLACKS	
	OLS	IV-2SLS	OLS	IV-2SLS	OLS	IV-2SLS
BOTH GENDERS						
All	0.073 (0.001)***	0.101 (0.005)***	0.077 (0.001)***	0.113 (0.006)***	0.053 (0.003)***	0.079 (0.012)***
HS Dropouts	0.047 (0.002)***	-0.031 (0.029)	0.054 (0.002)***	-0.014 (0.039)	0.027 (0.004)***	0.037 (0.025)
HS	0.059 (0.001)***	0.013 (0.017)	0.064 (0.002)***	0.028 (0.021)	0.038 (0.003)***	0.133 (0.015)***
Coll Dropouts	0.064 (0.001)***	0.055 (0.021)***	0.068 (0.001)***	0.064 (0.024)***	0.050 (0.003)***	0.086 (0.016)***
College	0.066 (0.001)***	0.040 (0.023)*	0.069 (0.001)***	0.054 (0.024)**	0.057 (0.004)***	0.102 (0.037)***
Graduate	0.065 (0.001)***	0.065 (0.019)***	0.067 (0.001)***	0.081 (0.023)***	0.071 (0.003)***	0.070 (0.033)**
MALES						
All	0.068 (0.001)***	0.084 (0.006)***	0.070 (0.001)***	0.098 (0.006)***	0.052 (0.005)***	0.053 (0.015)***
HS Dropouts	0.054 (0.002)***	-0.031 (0.034)	0.059 (0.003)***	0.011 (0.040)	0.036 (0.006)***	0.003 (0.035)
HS	0.058 (0.002)***	0.010 (0.020)	0.062 (0.002)***	0.036 (0.023)	0.039 (0.005)***	0.042 (0.014)***
Coll Dropouts	0.063 (0.002)***	0.043 (0.018)**	0.066 (0.002)***	0.066 (0.020)***	0.053 (0.005)***	0.111 (0.016)***
College	0.065 (0.002)***	0.042 (0.022)*	0.068 (0.002)***	0.063 (0.022)***	0.059 (0.007)***	0.108 (0.026)***
Graduate	0.063 (0.001)***	0.044 (0.019)**	0.065 (0.002)***	0.068 (0.022)***	0.070 (0.004)***	0.063 (0.039)*
FEMALES						
All	0.079 (0.001)***	0.116 (0.005)***	0.084 (0.001)***	0.121 (0.009)***	0.054 (0.003)***	0.108 (0.017)***
HS Dropouts	0.045 (0.003)***	0.012 (0.041)	0.052 (0.003)***	0.012 (0.055)	0.020 (0.005)***	0.045 (0.040)
HS	0.060 (0.002)***	0.039 (0.023)*	0.065 (0.002)***	0.045 (0.029)	0.035 (0.004)***	0.136 (0.022)***
Coll Dropouts	0.067 (0.002)***	0.074 (0.030)***	0.072 (0.002)***	0.073 (0.037)**	0.046 (0.004)***	0.073 (0.023)***
College	0.067 (0.002)***	0.085 (0.029)***	0.071 (0.002)***	0.082 (0.031)***	0.054 (0.005)***	0.189 (0.045)***
Graduate	0.069 (0.002)***	0.093 (0.028)***	0.071 (0.002)***	0.111 (0.035)***	0.072 (0.005)***	0.104 (0.032)***

Notes Instruments for education: CYS interacted with SOB dummies + continuing school dummy

*, ** and *** denote statistical significance at 90%, 95% and 99% confidence levels, respectively

Data is weighted by the person-specific sample weight

White's robust standard errors are reported in parentheses

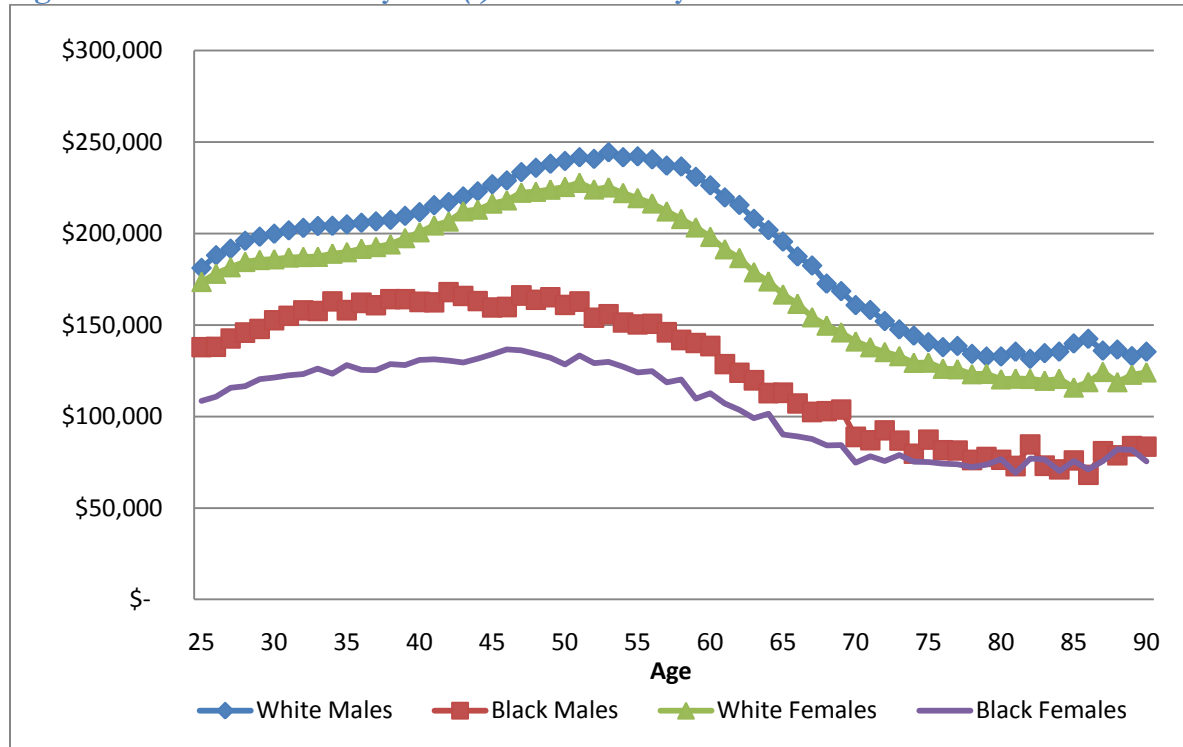
$$R(E) = R = \delta_1: \quad \ln(y_i) = \delta_0 + \delta_1 E_i + \delta_2' X_i + \varepsilon_i \quad (20)$$

$$R(E \in l) = \pi^l: \quad \ln(y_i) = \pi_0 + \sum_{l=1}^5 \pi^l E_i * EL_i^l + \pi_2' X_i + \omega_i \quad (21)$$

The sample is comprised of non-institutionalized US natives (excluding natives or residents of Alaska and Hawaii) aged 51-82. Monetary returns are based on a Mincerian regressions of the natural logarithm of per capita household income on the following controls: years of completed schooling, experience, experience squared, and dummies for: state of residence, marital status (single, widowed, and divorced; married is the reference category), urban residence, employment status (unemployed, disabled, and out of the labor force; employed is the reference category), and inter-state migration. To allow returns to vary across schooling levels, E_i is interacted with education level dummies EL_i^l which equal one if individual i is at level l . The instruments for education are the level of CYS interacted with SOB and the continuation school dummy. OLS and IV-2SLS estimates use person-specific sample weights and White's robust standard errors.

Figures

Figure 1. The value of a life year $v(t)$ over the lifecycle

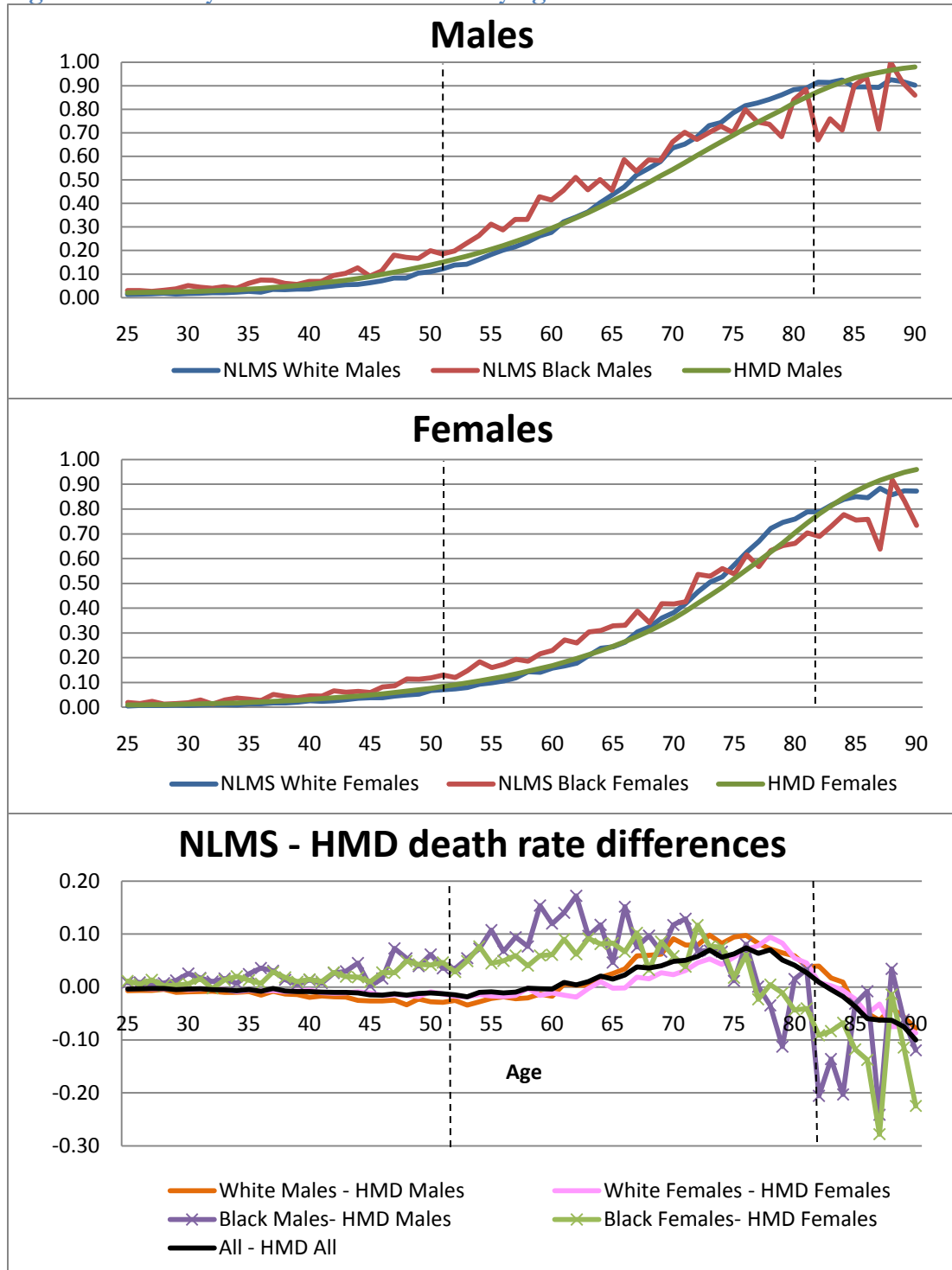


The value of a life-year is estimated with the equation:

$$v(t) = w(t; E)[1 + \Phi(z)] \quad (14)$$

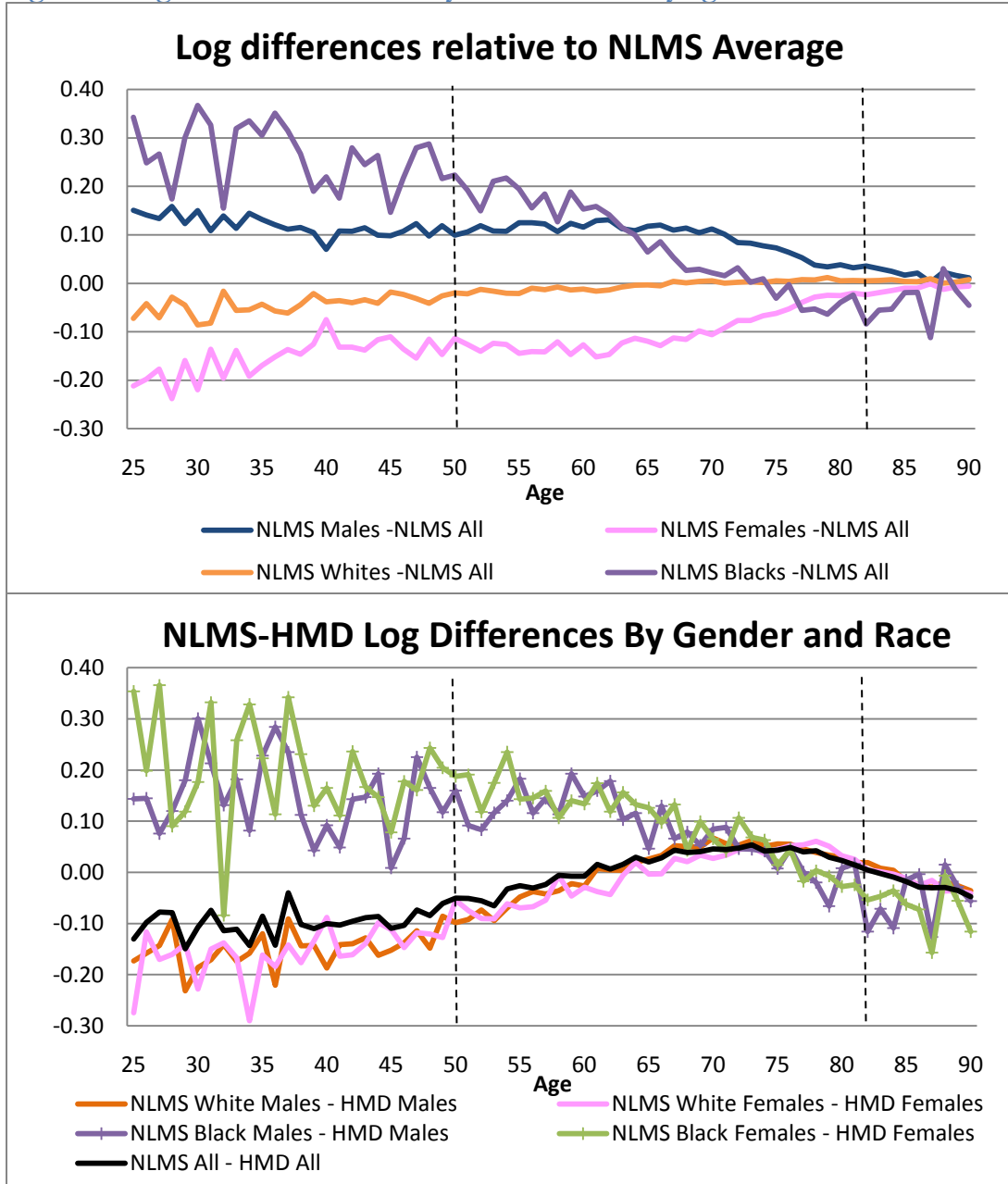
Wage earnings $w(t; E)$ are thrice the value of per capita household income smoothed over the lifecycle. Consumer surplus $\Phi(z)$ is calibrated at 0.64 to yield a statistical value of life (SVL) of \$5 million across the entire sample with an interest rate of 3%. Dollar values are expressed in 2004 USD.

Figure 2. Eleven-year death rates in 1983 by age: NLMS vs. HMD



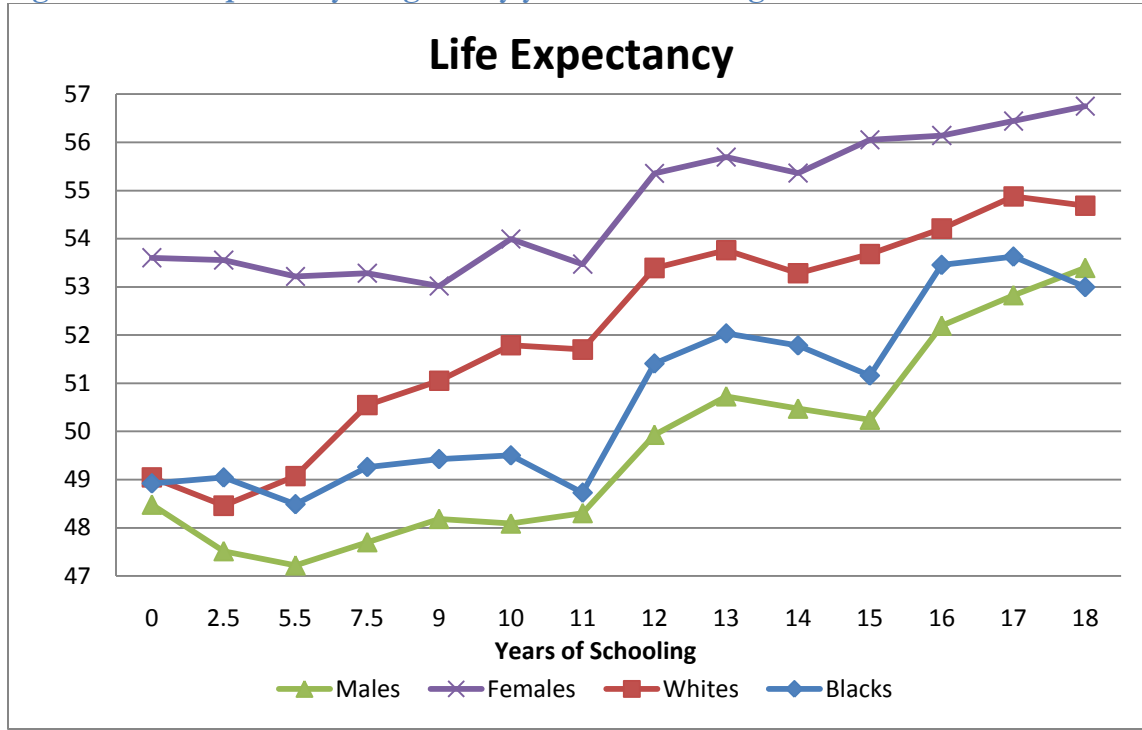
HMD annual death rates at age t ($\lambda_1(t)$) were converted to eleven-year mortality measures $\lambda_{11}(t)$ with the formula: $\lambda_{11}(t) \cong 1 - e^{-\sum_{\tau=0}^{10} \lambda_1(t+\tau)}$.

Figure 3. Log differences in eleven-year death rates by age: NLMS vs. HMD



NLMS mortality rates, unlike Human Mortality Database (HMD) estimates, exclude the institutionalized US population causing eleven-year death rates of blacks beyond age 75 to be underestimated. HMD annual death rates at age t ($\lambda_1(t)$) are converted to eleven-year mortality measures $\lambda_{11}(t)$ with the formula: $\lambda_{11}(t) \cong 1 - e^{-\sum_{\tau=0}^{10} \lambda_1(t+\tau)}$.

Figure 4. Life expectancy at age 25 by years of schooling

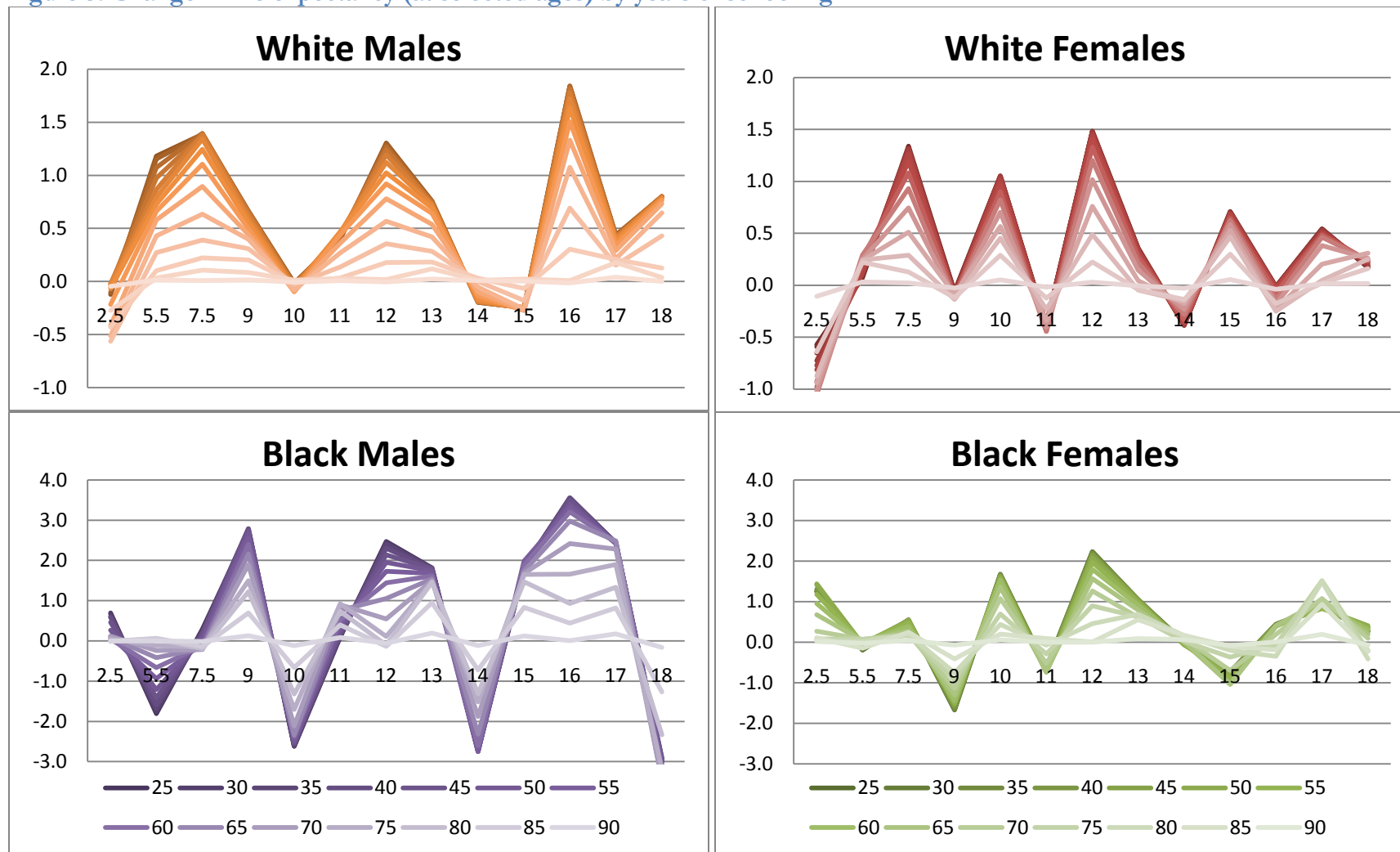


Life expectancy at age 25 for a person with E years of schooling is computed as:

$$LE(a = 25; E) = \int_a^\infty \tilde{S}(t, a; E) dt \quad (15)$$

The survival probability from age a to age t is defined in terms of the annual hazard rate λ_1 , or one-year probability of death conditional on reaching age τ : $\tilde{S}(t, a; E) \equiv \prod_{\tau=a}^t (1 - \lambda_1(\tau; E))$.

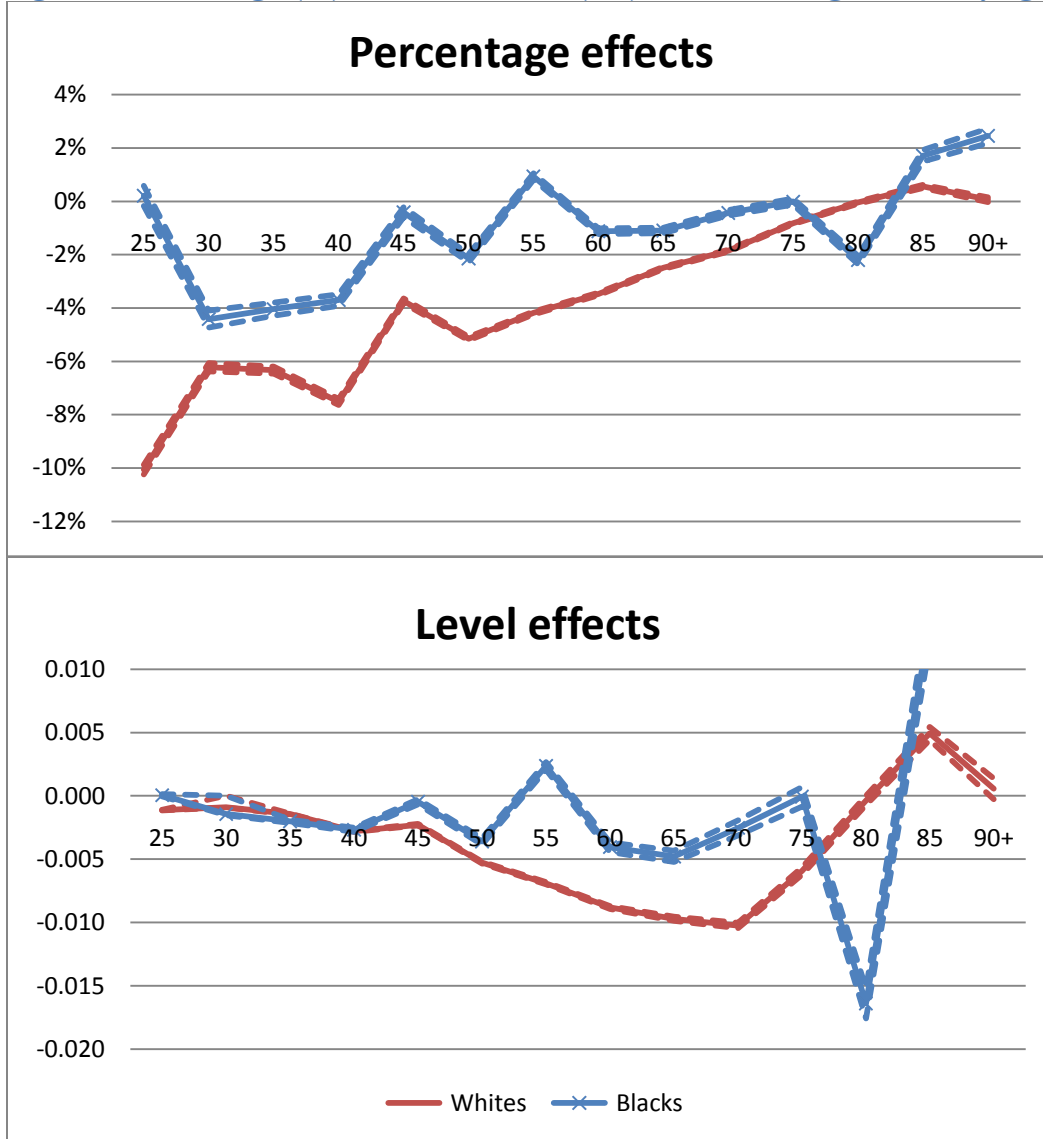
Figure 5. Change in life expectancy (at selected ages) by years of schooling



The change in life expectancy at age a from an additional year of schooling at level E is computed as:

$$\Delta LE(a; E) \equiv LE(a; E + 1) - LE(a; E) = \int_a^\infty [\tilde{S}(t, a; E + 1) - \tilde{S}(t, a; E)] dt$$

Figure 6. Percentage (ψ_1) and level effects ($\psi_1\lambda$) of Cox PH regressions by age group

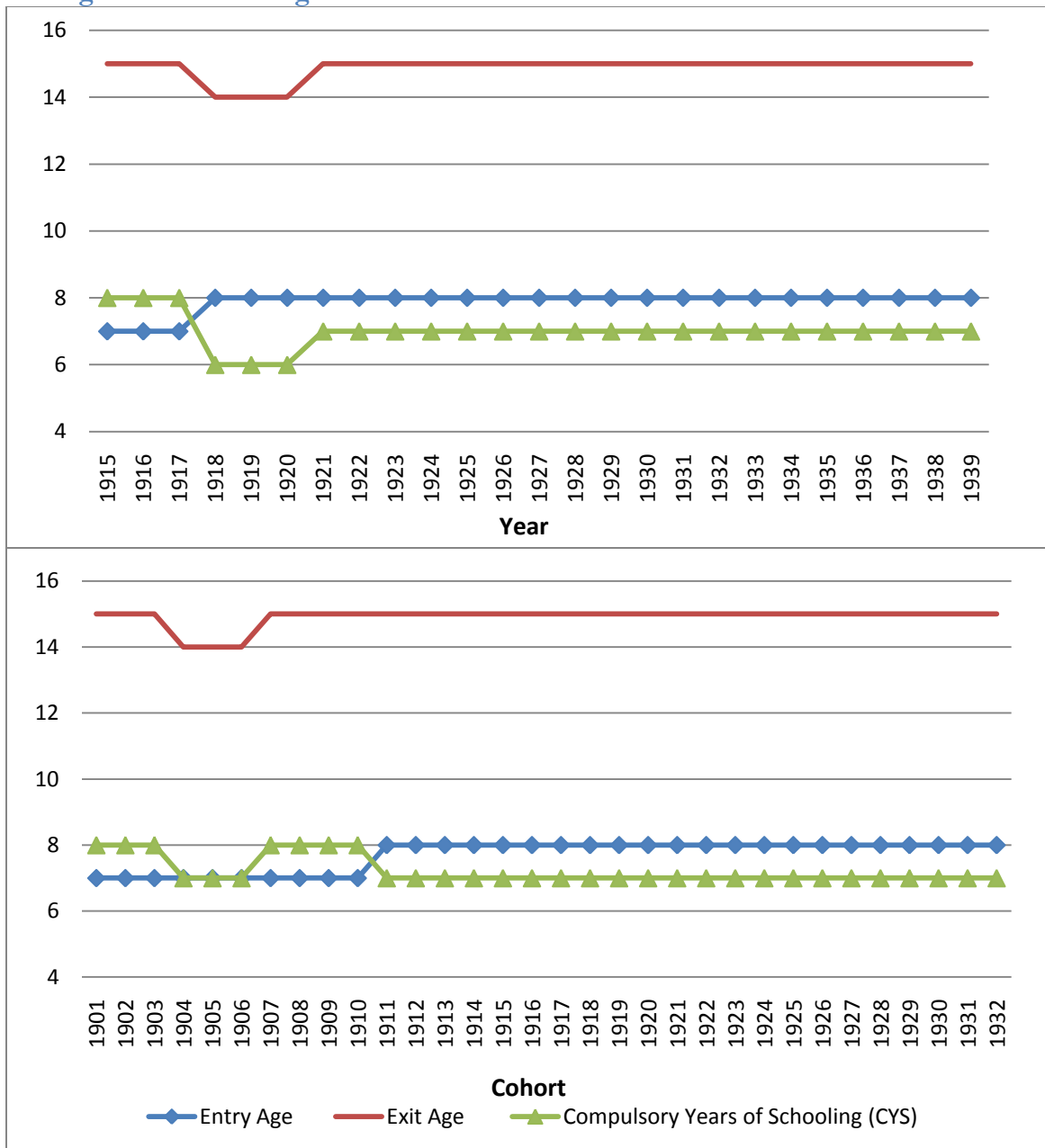


The percentage (ψ_1) and level effects ($\psi_1\lambda$) of an additional year of schooling on 11-year mortality are based on the following Cox PH model run separately on 5-year age groups:

$$\lambda_i(t = 11; E_i, X_i) = \theta_t e^{\psi_1 E_i + \psi_2' X_i} \quad (\text{Cox PH})$$

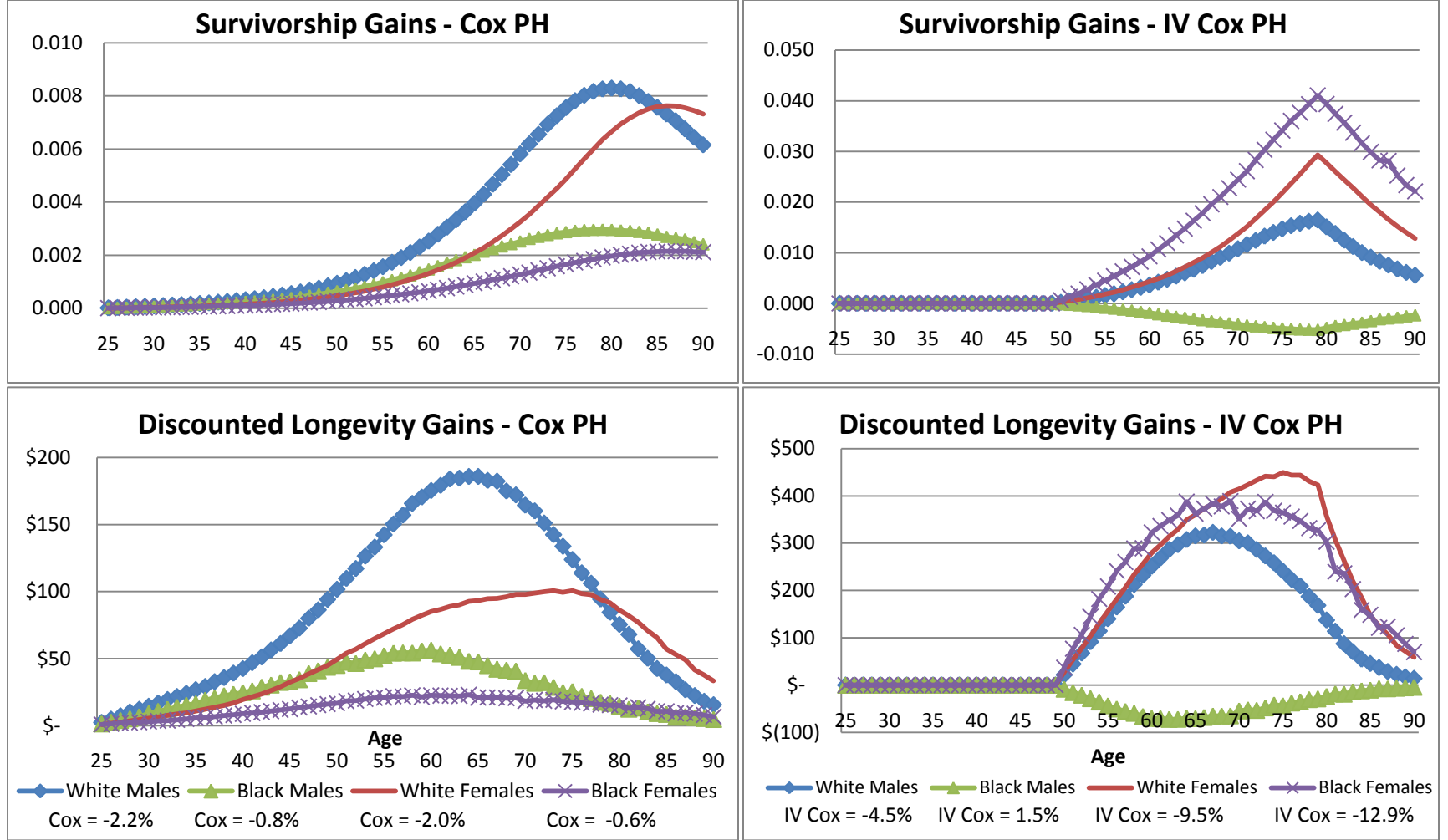
$\lambda_i(t = 11; E, X)$ is the 11-year hazard rate (with time-dependent hazard factor θ_t) for an individual i with E_i years of completed education and X_i denotes a set of controls including the natural logarithm of per capita income and dummies for: female gender, state of residence, marital status, urban residence and employment status. Data is weighted by individual sampling weights and the dotted lines represent the estimates' 95% confidence intervals.

Figure 7. Translating CSL into CYS: the case of California.



Entry age is the maximum age by which children must enter school. Exit age is the minimum age at which they may leave school (usually the earliest age at which they can get a work permit). Compulsory years of schooling (CYS) is the difference between exit age and entry age.

Figure 8. Survivorship gains (top) and monetized longevity gains of schooling (bottom): Cox PH and IV Cox PH models



Survivorship gains are:

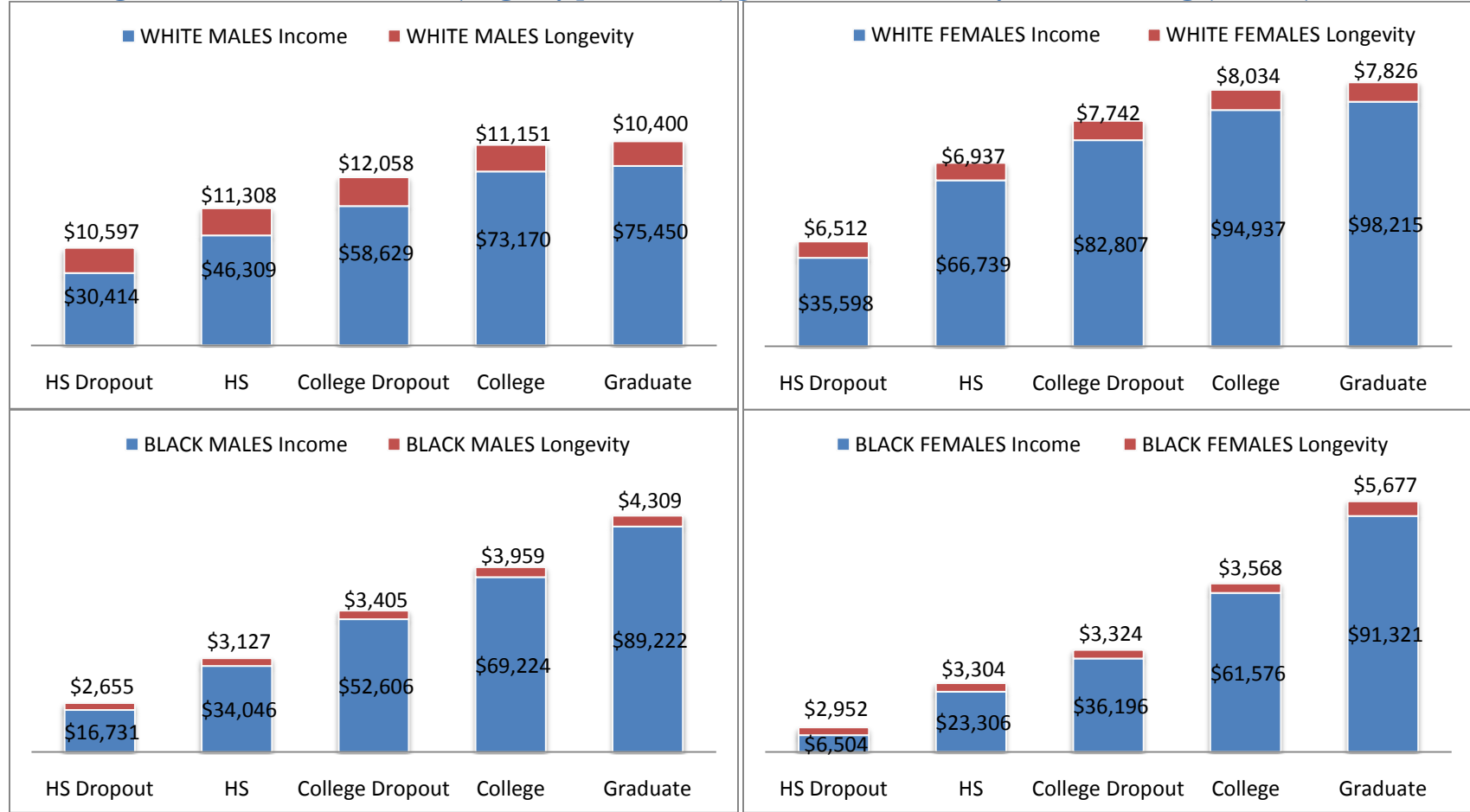
$$\frac{\partial \tilde{S}(t,a;E)}{\partial E} \cong \psi_1 \prod_{\tau=a}^t \lambda_1(\tau; E) \quad (18)'$$

Average discounted longevity gains at age t are:

$$v(t) \frac{\partial \tilde{S}(t,a;E)}{\partial E} e^{-r(t-a)} \tilde{S}(t, a = 25; E); r = 3\%.$$

Cox PH (ψ_1) and IV Cox PH (κ_1) marginal effects from Tables 4 and 7 are reported in percentage terms.

Figure 9. Present value of total (longevity plus income) gains of an additional year of schooling ($r = 3\%$)

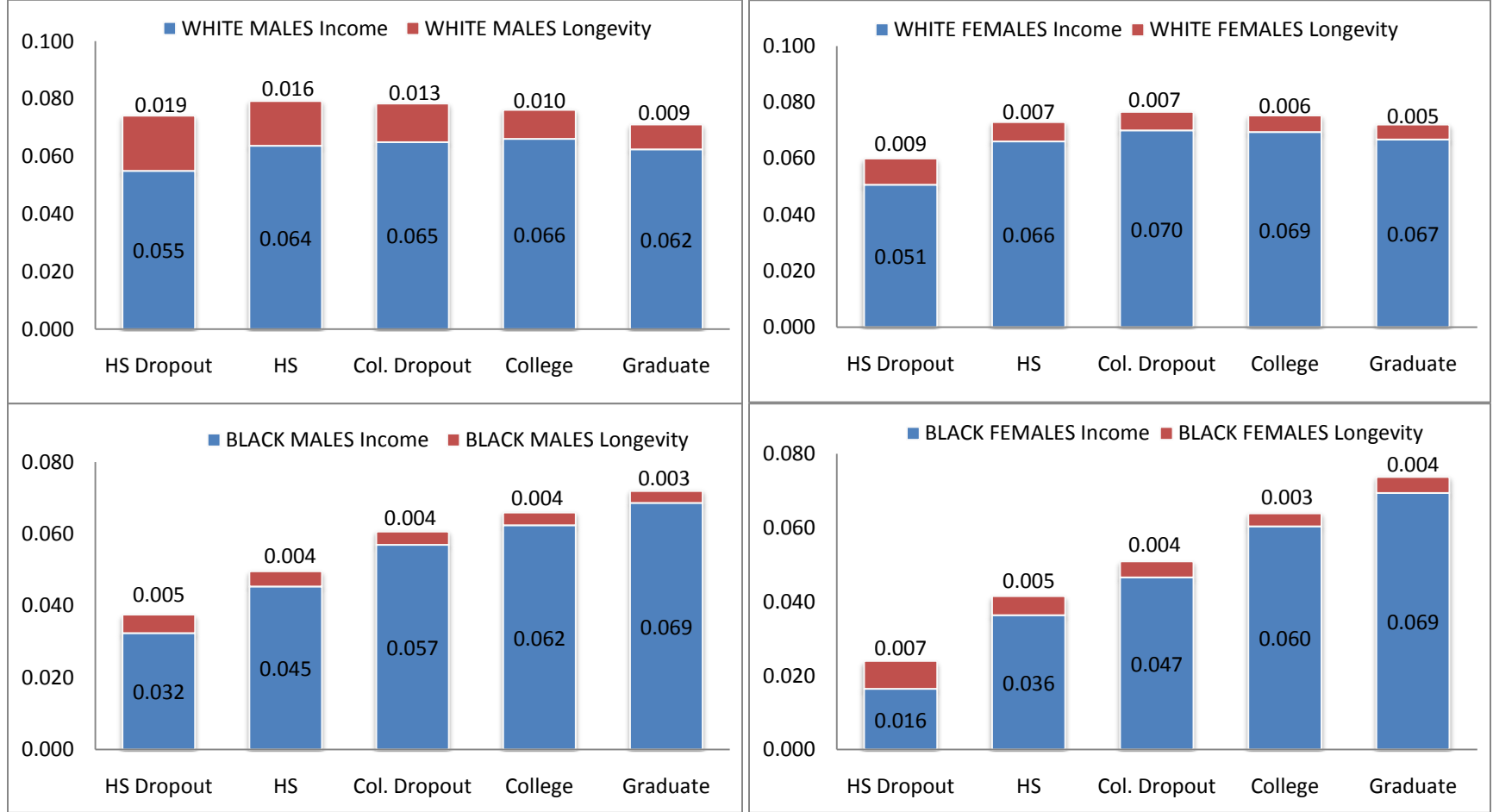


The total gains of an additional year of education are based on equation (25):

$$V_T(a = 25; E) = \int_a^\infty \left\{ v(t) \frac{\partial \tilde{S}(t, a; E)}{\partial E} + \frac{\partial y(t; E)}{\partial E} \right\} e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (8)$$

$$V_T(a = 25; E) \cong \int_a^\infty \{ [\tilde{S}(t, a; E + 1) - \tilde{S}(t, a; E)] v(t) + R(E) y(t; E) \} e^{-r(t-a)} \tilde{S}(t, a; E) dt \quad (25)$$

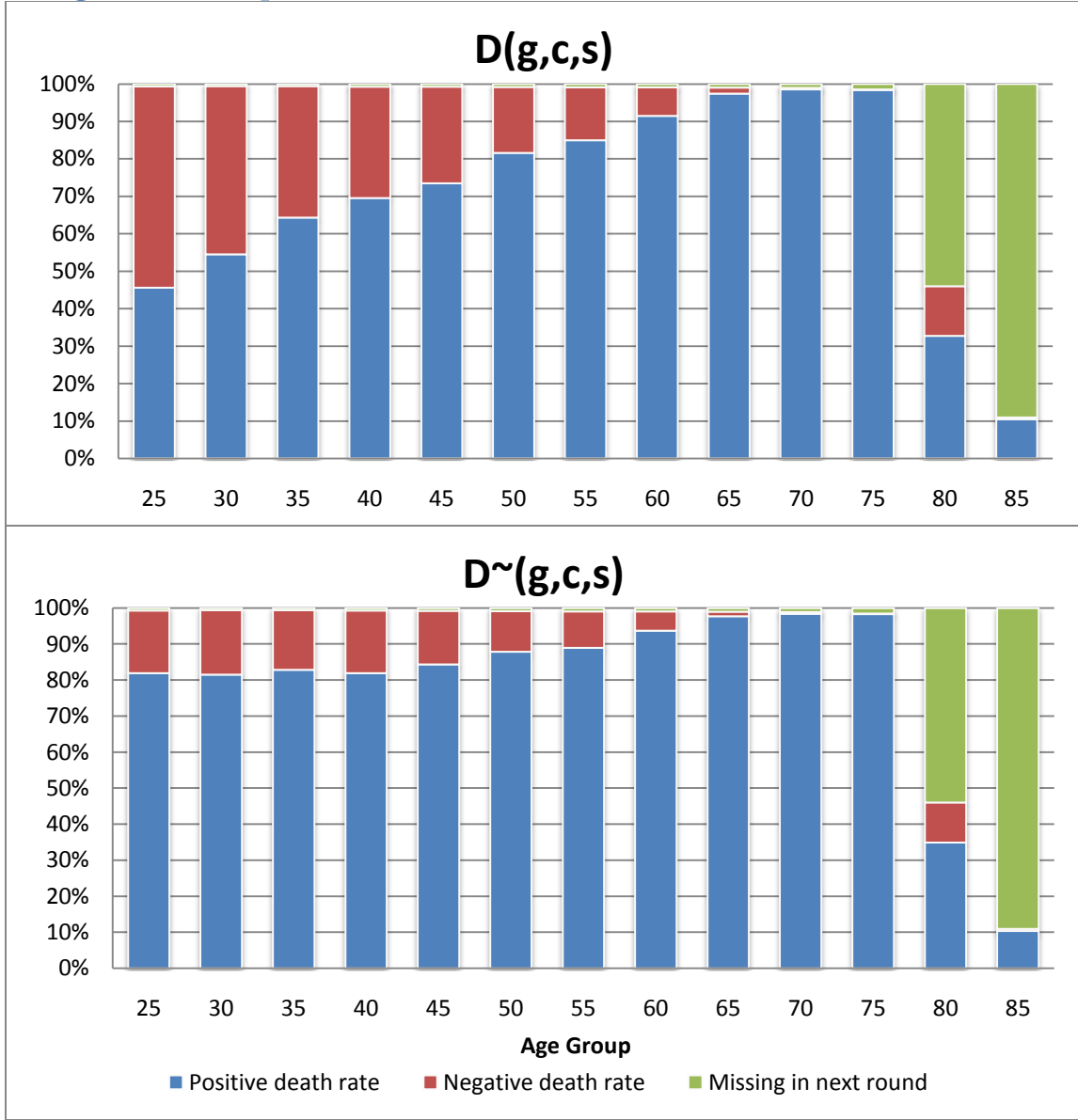
Figure 10. Total (longevity θ plus income R) returns to schooling ($r = 3\%$)



Total returns to an additional year of schooling $\Theta(E)$ are the sum of income returns $R(E)$ (OLS Mincerian estimates by education level) and longevity returns $\theta(E) \equiv R(E) \frac{V_L(E)}{V_Y(E)}$:

$$\Theta(E) \equiv \theta(E) + R(E) = R(E) \left(\frac{V_T(E)}{V_Y(E)} \right) \quad (27)$$

Figure 11. Description of IPUMS death rates



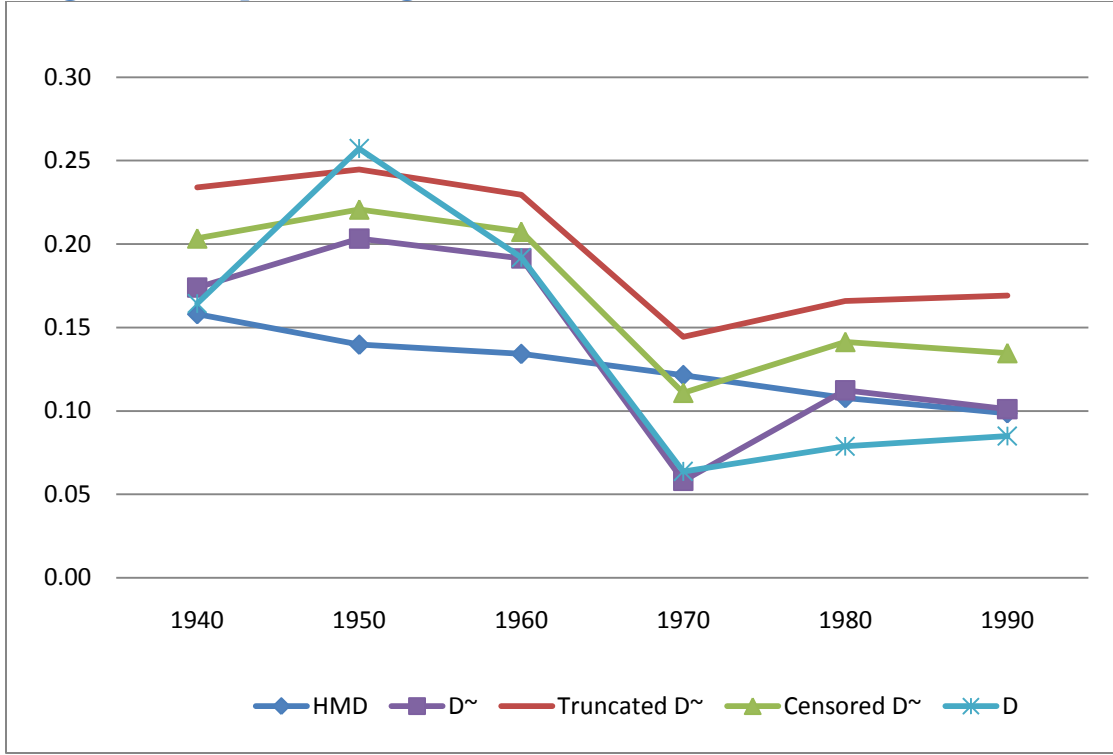
The ten-year death rates of group (g, c, s) (gender g of cohort c born in state s) with size $N_t(g, c, s)$ at period $t = 1940, \dots, 1990$ are:

$$D_t(g, c, s) \equiv \frac{N_t(g, c, s) - N_{t+10}(g, c, s)}{N_t(g, c, s)} = \frac{\sum_{i=N_{t+10}(g, c, s)}^{N_t(g, c, s)} 1}{\sum_{i=1}^{N_t(g, c, s)} 1} \quad (19)$$

$$\tilde{D}_t(g, c, s) \equiv \frac{\tilde{N}_t(g, c, s) - \tilde{N}_{t+10}(g, c, s)}{\tilde{N}_t(g, c, s)} = \frac{\sum_{i=N_{t+10}(g, c, s)}^{N_t(g, c, s)} \omega_i^t(g, c, s)}{\sum_{i=1}^{N_t(g, c, s)} \omega_i^t(g, c, s)} \quad (20)$$

IPUMS data consists of 1% random samples from the 1940-2000 decennial Census rounds. Including only US natives (neither born nor residing in Alaska or Hawaii) 25 years or older yields a sample size of 7,258,182.

Figure 12. Comparison of age-standardized death rates: IPUMS vs HMD



Ten-year death rates are defined as follows:

$HMD \equiv$ ten-year mortality rates from Human Mortality Database

$$D \equiv D_t(g, c, s) \equiv \frac{N_t(g, c, s) - N_{t+10}(g, c, s)}{N_t(g, c, s)} = \frac{\sum_{i=N_{t+10}(g, c, s)}^{N_t(g, c, s)} 1}{\sum_{i=1}^{N_t(g, c, s)} 1} \quad (28)$$

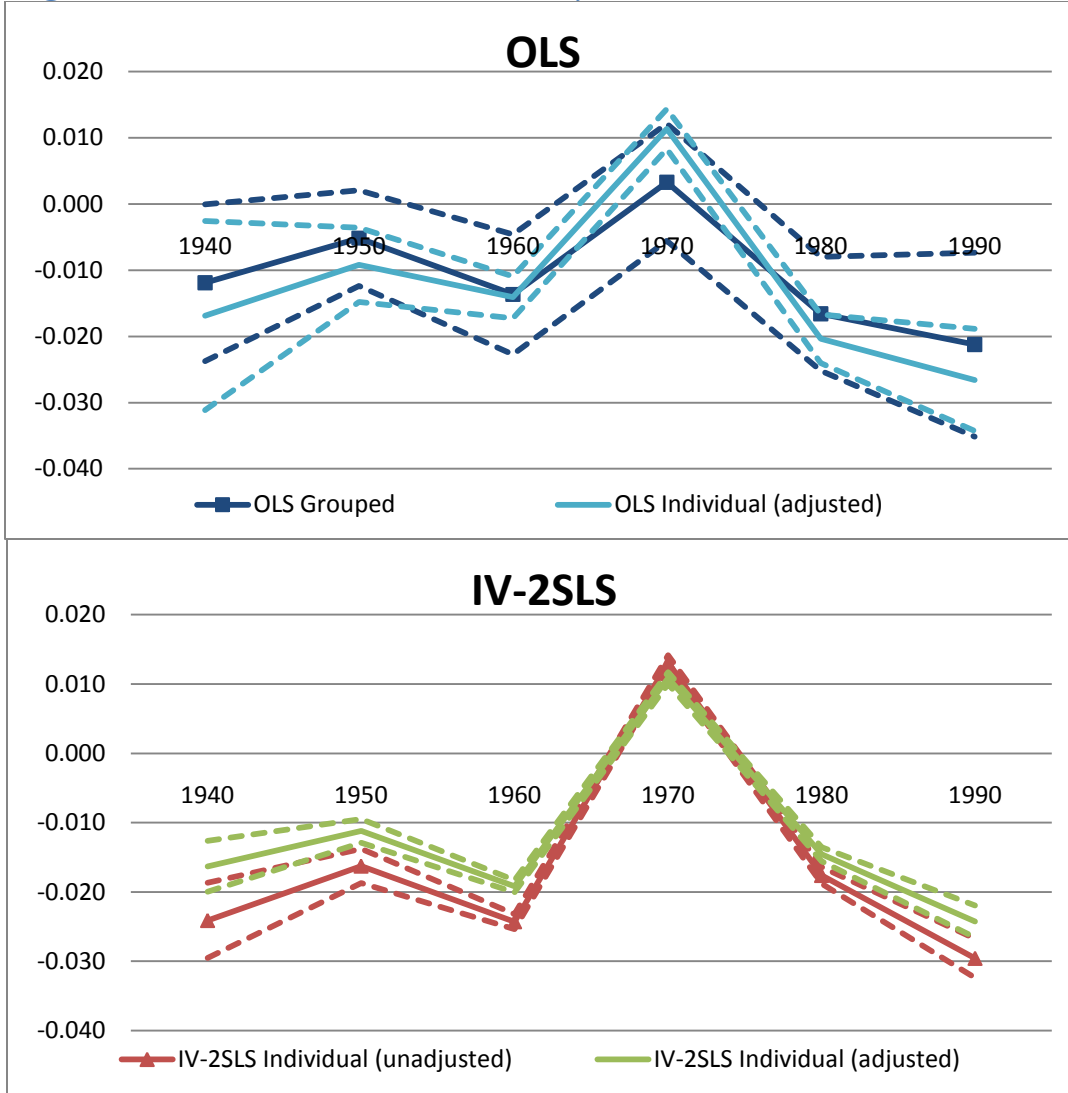
$$D \sim \equiv \tilde{D}_t(g, c, s) \equiv \frac{\tilde{N}_t(g, c, s) - \tilde{N}_{t+10}(g, c, s)}{\tilde{N}_t(g, c, s)} = \frac{\sum_{i=N_{t+10}(g, c, s)}^{N_t(g, c, s)} \omega_i^t(g, c, s)}{\sum_{i=1}^{N_t(g, c, s)} \omega_i^t(g, c, s)} \quad (29)$$

$Censored D \sim \equiv D \sim$ with negative values set to zero

$Truncated D \sim \equiv D \sim$ with negative values dropped

The term $\tilde{N}_t(g, c, s) \equiv \sum_{i=1}^{N_t(g, c, s)} \omega_i^t(g, c, s)$ is the sum of individual sampling weights $\omega_i^t(g, c, s)$ over the $N_t(g, c, s)$ members of group (g, c, s) in period t . All rates are age-standardized using population shares from 1990.

Figure 13. OLS and IV-2SLS estimates of β ; 1940-1990



Dotted lines represent 90% confidence intervals. The ten-year death rates is calculates as:

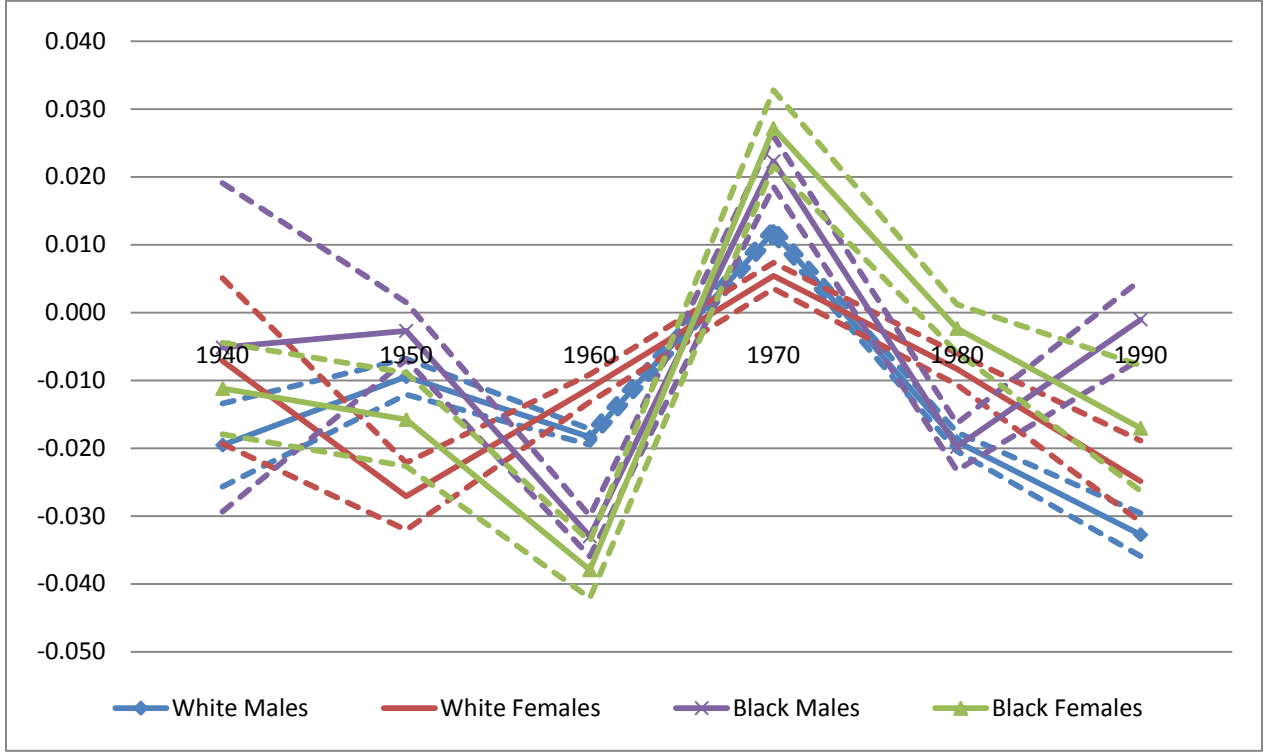
$$\tilde{D}_t(g, c, s) \equiv \frac{\tilde{N}_t(g, c, s) - \tilde{N}_{t+10}(g, c, s)}{\tilde{N}_t(g, c, s)}, \text{ where } \tilde{N}_t(g, c, s) \equiv \sum_{i=1}^{N_t(g, c, s)} \omega_i^t(g, c, s) \text{ and } \omega_i^t(g, c, s) \text{ is the}$$

individual sampling weight. Controls include: five-year age group, female, age group interacted with female, state of residence, marital status and employment status dummies. In IV-2SLS regressions, five-year age group dummies are replaced by 1901-1932 cohort dummies. The four estimates presented are: i) $\hat{\gamma}_1^{OLS} = OLS \text{ Grouped}$; ii) $\hat{\alpha}_1^{ADJ} = OLS \text{ Individual (adjusted)}$; iii) $\hat{\alpha}_1^{IV} = IV\text{-}2SLS \text{ Individual (unadjusted)}$; and iv) $\hat{\alpha}_1^{IV-ADJ} = IV\text{-}2SLS \text{ Individual (adjusted)}$ based on the following models:

$$\tilde{D}_t(g, c, s) = \alpha_0 + \alpha_1 E_{it}(g, c, s) + \alpha_2 X_{it}(g, c, s) + \varepsilon_{it}(g, c, s) \quad (30)$$

$$\tilde{D}_t(g, c, s) = \gamma_0 + \gamma_1 \bar{E}_t(g, c, s) + \gamma_2 \bar{X}_t(g, c, s) + \bar{\varepsilon}_t(g, c, s) \quad (31)$$

Figure 14. IV-2SLS estimates ($\hat{\alpha}_1^{IV}$) by gender and race: 1940-1990

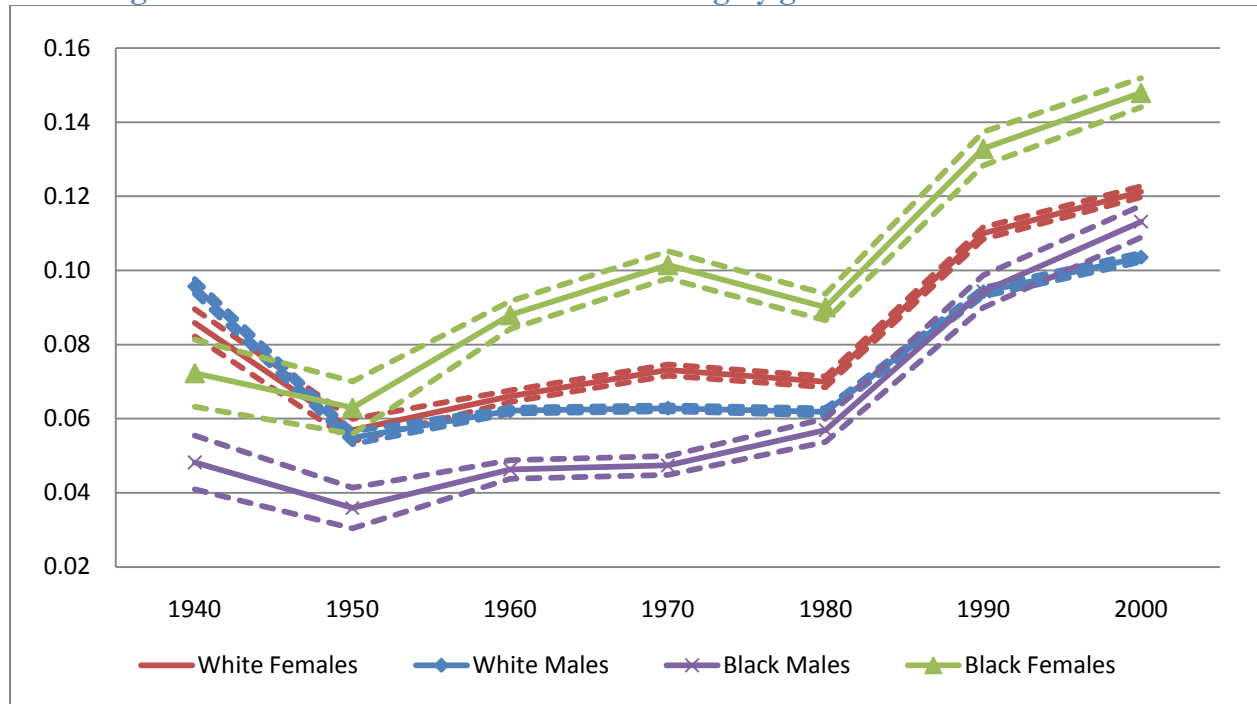


Dotted lines represent 90% confidence intervals. The ten-year death rates is calculated as: $\tilde{D}_t(g, c, s) \equiv \frac{\tilde{N}_t(g, c, s) - \tilde{N}_{t+10}(g, c, s)}{\tilde{N}_t(g, c, s)}$, where $\tilde{N}_t(g, c, s) \equiv \sum_{i=1}^{N_t(g, c, s)} \omega_i^t(g, c, s)$ and $\omega_i^t(g, c, s)$ is the individual sampling weight. IV-2SLS estimate $\hat{\alpha}_1^{IV}$ is based on the following regression carried out separately by gender, race and Census round t :

$$\tilde{D}_t(g, c, s) = \alpha_0 + \alpha_1 E_{it}(g, c, s) + \alpha_2' X_{it}(g, c, s) + \varepsilon_{it}(g, c, s) \quad (30)$$

Controls X_{it} include dummies for: 1901-1932 cohorts, state of residence, marital status and employment status. All regressions use sample-line weights and White's robust standard errors.

Figure 15. OLS Mincerian returns to schooling by gender and race: 1940-2000



Dotted lines represent 90% confidence intervals. The dependent variable ($\ln(y_i)$) is the natural logarithm of annual wage earnings. Controls X_i include: years of *age*, *age* squared, and dummies for: SOB, state of residence, marital status, and employment status. Regressions are carried out separately by gender, race and Census round for individuals 25 years or older using person-specific sample-line weights and White's robust standard errors.

$$\ln(y_i) = \delta_0 + RE_i + \delta_X'X_i + \varepsilon_i \quad (\text{Mincerian regression})$$