

Henderson, Daniel J.; Polachek, Solomon; Wang, Le

Working Paper

Heterogeneity in schooling rates of return

IZA Discussion Papers, No. 5662

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Henderson, Daniel J.; Polachek, Solomon; Wang, Le (2011) : Heterogeneity in schooling rates of return, IZA Discussion Papers, No. 5662, Institute for the Study of Labor (IZA), Bonn,
<https://nbn-resolving.de/urn:nbn:de:101:1-201104283170>

This Version is available at:

<https://hdl.handle.net/10419/51822>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 5662

Heterogeneity in Schooling Rates of Return

Daniel J. Henderson
Solomon W. Polachek
Le Wang

April 2011

Heterogeneity in Schooling Rates of Return

Daniel J. Henderson

*State University of New York at Binghamton
and IZA*

Solomon W. Polachek

*State University of New York at Binghamton
and IZA*

Le Wang

University of New Hampshire

Discussion Paper No. 5662

April 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Heterogeneity in Schooling Rates of Return^{*}

This paper relaxes the assumption of homogeneous rates of return to schooling by employing nonparametric kernel regression. This approach allows us to examine the differences in rates of return to education both across and within groups. Similar to previous studies we find that on average blacks have higher returns to education than whites, natives have higher returns than immigrants and younger workers have higher returns than older workers. Contrary to previous studies we find that the average gap of the rate of return between white and black workers is larger than previously thought and the gap is smaller between immigrants and natives. We also uncover significant heterogeneity, the extent of which differs both across and within groups. The estimated densities of returns vary across groups and time periods and are often skewed. For example, during the period 1950-1990, at least 5% of whites have negative returns. Finally, we uncover the characteristics common amongst those with the smallest and largest returns to education. For example, we find that immigrants, aged 50-59, are most likely to have rates of return in the bottom 5% of the population.

JEL Classification: C14, J24

Keywords: Mincer regressions, nonparametric, rate of return to education

Corresponding author:

Daniel J. Henderson
Department of Economics
State University of New York
Binghamton, NY 13902-6000
USA
E-mail: djhender@binghamton.edu

^{*} The authors would like to thank participants of the 2009 SOLE meeting in Boston, MA as well as participants of the 2009 IZA/SOLE Transatlantic Meeting of Labor Economists (especially Kasey Buckles, Dan Hamermesh, Stephen Trejo and Ian Walker) in Ammersee, Germany, for useful comments which led to an improved version of this paper. The research on this project was started when Wang was at the Minnesota Population Center at the University of Minnesota and he would like to thank them for their data and computer support.

1 Motivation

Much economic research analyzes the behavior of a typical economic agent, but ignores the possibility that agents differ from each other. This assumption is prevalent in labor economics but is often employed in other areas of economics. For example, in microeconomics, theoreticians often assume a given utility or production function but ignore heterogeneous agents. As an illustration, Jaimovich and Rebelo (2009) state “our model economy is populated by *identical* (emphasis ours) agents who maximize their lifetime utility . . . [and] output is produced with a Cobb-Douglas production function” In macroeconomics, practitioners estimate the aggregate effect of a change in monetary policy but often downplay the fact that various agents and economic sectors can react differently. For example, the dynamic stochastic generalized equilibrium model describes the economy in terms of a three equation model in which one equation is an interest rate feedback mechanism used by the central bank, a second is an Euler consumption equation, and a third is a price setting scheme aggregated from monopolistically competitive firms, but in which the heterogeneous agents are not treated independently (Christiano, Eichenbaum and Evans, 2005; Smets and Wouters, 2003). In the subfields, for example in public finance, Parry and Small (2009) derive “tractable formulas for the welfare effects of fare adjustments in passenger peak and off-peak rail and bus transit” based on a representative agent framework, but recognize “that this can only approximate the aggregate behavior of a diverse population” (pp. 273).

But of late, the advantages of examining population-wide heterogeneity are becoming more widespread. For example, with regard to the debate whether prices are flexible or sticky, Boivin, Giannoni and Mihov (2009) allude to the fact that “empirical studies based on aggregate data . . . found stickiness” whereas “evidence based on the behavior of disaggregated prices suggests that prices are much more volatile than conventionally assumed in studies based on aggregate data” (pp. 350-351). In reconciling the difference, they then evaluate different responses to macroeconomic and sector-specific shocks. Similarly in economic theory, models are beginning to address heterogeneity in a variety of ways. For example, Chiappori and Ekeland (2009) analyze “when, to what extent and under what conditions” one can recover underlying individual preferences and decision processes from a group’s aggregate behavior.

Similarly, most econometric applications neglect individual differences by concentrating on population-wide estimates. However, at least since 1950, econometricians considered the case when coefficients vary across individual observations (Rubin, 1950). Early empirical

work consisted of random coefficients models. The problem, however, is that these type models impose strict distributional restrictions (usually normality) on the heterogeneous coefficients, and anyway end up estimating the mean response, as in Hildreth and Houck (1968). Further, the approach assumes a particular functional form for the underlying equation. This is also true of extensions of the random coefficients model, for example, the correlated random coefficients model (Heckman, Schmieder and Urzua, 2009).

Later work used panel data with fixed-effects methods to adjust for heterogeneity. Of these, most analyses assume person specific intercepts. Person specific intercepts mean individuals differ only with respect to the level of the outcome variable, but such models do not allow exogenous variables to affect each individual differently. But clearly individuals can vary in the ways they react to exogenous forces. For example, one can conceive an extra year of schooling might benefit a high ability person differently than a low ability person. Models that adopt individual specific slopes generally do not consider more than one exogenous variable. But even so, they still estimate single common population-wide parameters for each of the other independent variables (Polachek and Kim, 1994; Pesaran, 2006). But more crucially not always are panel data available to utilize fixed or random-effects techniques; and even if they are, it is not clear one should concentrate on a single summary parameter for the population rather than a distribution of parameters across individuals within the population.

In this paper we adopt nonparametric kernel estimation. This approach does not necessarily take the place of panel data when such data are available. However, one advantage is such estimation techniques need not rely on panel data which frequently are not available. Also, nonparametric models do not require the functional form to be specified *a priori*. We apply the approach to get at the distribution of educational rates of return. Knowing the extent individual-specific rates of return vary is important to policy makers, not only because it gives an indication of the benefits of schooling individuals accrue, but also because rates of return have implications regarding race differences in earnings (Card and Krueger, 1992), because rates of return give an indication of supply and demand shifts in the labor market (Freeman, 1977; Card and Lemieux, 2001), and because rates of return have implications for technological change (Goldin and Katz, 2008). Thus our purpose is to document heterogeneity in a systematic and coherent manner.

There are a number of ways to estimate rates of return. Early studies, namely Becker (1964), computed the net benefits of education that weighed the gains in earnings from schooling against tuition (direct costs) and opportunity costs. Heckman, Lochner and Todd

(2006) advocate a computation involving option values. However, in 1974, the Mincer earnings function approach became the norm. Today literally hundreds of studies adopt this approach (Polachek, 2008).

The major problem with all these rate of return studies is the difficulty in getting at heterogeneity. To adopt fixed-effects one needs panel data, but within the panel each individual must exhibit some variation in schooling, i.e., a set of individuals must increase their schooling levels at various points in their working lives. The problem, though, is that those working during school have arbitrarily low earnings (Lazear, 1977), typically because of time and geographic constraints limit one's ability to commute. As noted in Card (1995), these individuals appear to have larger returns to education than others, and thus panel data estimation may over-estimate the average returns to education. But even if not, it is difficult to isolate increases in earnings caused by schooling changes from increases in earnings caused by increases in experience as individuals age. As such, there is very little research trying to account for person-specific heterogeneity in schooling rates of return, and those that do rely on distributional assumptions (Harmon, Hogan and Walker, 2003) or priors (Koop and Tobias, 2004).¹ For this reason we adopt nonparametric kernel estimation which does not rely on restrictive functional form assumptions or on panel data.

To anchor our results to these past studies, we choose the Mincer notion of rates of return rather than other notions such as Becker's that uses both direct and opportunity costs or Heckman, Lochner and Todd's (2006) that use option values since, as we indicated, most current rate of return estimates have adopted Mincer's earnings function approach we suspect because of its tractability; but obviously other rate of return definitions could have been employed. However, using other definitions would yield too few studies to compare to our new nonparametric kernel estimates. For this reason we adopt the Mincer specification, and this is why we make great efforts to compare our nonparametric estimates to typical parametric results.

The rest of the paper is organized as follows. Section 2 introduces the overall approach and previews some important results; Section 3 presents the empirical method in detail; Section 4 describes the data; Section 5 discusses the results; Section 6 identifies characteristics of those with the highest and lowest returns; Section 7 explains the implications of our results

¹ Recently, there have also been studies utilizing quantile regression methods (see, e.g., Martins and Pereira, 2004). Quantile regression methods typically focus on a specific heterogeneity effects across the whole earnings distribution. Controlling for a set of covariates in estimation – conditional quantile regression – allows for the interpretation of within-group effects. However, we shall show, there exists a larger heterogeneity in returns to schooling, both between- and within-groups.

for future research; Section 8 concludes.

2 Overall Approach

The rate of return to schooling literature is not ignorant of these econometric developments regarding heterogeneity (Koop and Tobias, 2004), however, most of the literature assumes that all individuals within a particular group (for example, white males) have the same return to a one-unit increase in education.² There are good reasons to believe that rate of return to schooling varies within a particular sub-group. For example, for white males some individuals have higher ability, some have better access to credit, some attend higher quality schools, some specialize in market oriented subjects and some obtained schooling more recently. It is also natural to believe that this variation may differ across groups (such as between blacks and whites) and over time (such as each decade from 1940 to 2000). Uncovering this variability may contain useful information for policy makers. Panel data estimators relax the assumption of identical rates of return for each individual, but they require repeated observations in which individuals increase their education during their working lives. This information is not always readily available. However, even if it were, arbitrarily lower earnings before finally terminating school leads to overestimates of rates of return (Card, 1995). Also, the level of education for a particular individual is generally constant later in life and hence would be removed by typical fixed effects estimators. Random coefficients models allow for parameter heterogeneity and estimation in a cross-section, but are subject to the same functional form restrictions as typical panel data estimators (Harmon, Hogan and Walker, 2003). If the specified parametric functional form is incorrect, estimation generally leads to inconsistent estimates. Further, these methods often require restrictive assumptions regarding the variability (for example, symmetric distributions). On the other hand, non-parametric regression methods allow for separate rate of return estimates with respect to each realization of the regressors. They also eschew functional form assumptions and are consistent under a more broad range of data generating processes (Heckman, Lochner and Todd, 2008). In this paper we use both parametric and nonparametric methods to investigate heterogeneity across groups. We further use the nonparametric methods to investigate heterogeneity within groups (for example, to plot out the distribution of rates of return for white males and black males by year).

² Although it can be argued that schooling and education are distinct, we speak of them as being synonymous in this paper.

To preview our results, we first note that our parametric results are generally consistent with the literature. We find that the rate of return has been rising over time, albeit non-monotonically. The rate of return to older individuals is less than that of younger workers. The return black workers receive is greater than their white counterparts. We also find that the rate of return is higher for natives than immigrants. Following cohorts over time we find that the rate of return falls with age, but rises within a particular age-group over time. We also find that the drop in the rate of return over time is faster for blacks than for whites. A similar result is found for native versus immigrant workers.

The results from the nonparametric regressions on average are often in line with the parametric findings. We find that the rate of return to education (on average) is larger for younger as opposed to older individuals. We also find that blacks have a higher rate of return than whites and that natives have higher rates of return of schooling than immigrants. However, we find that the absolute difference between the returns of each set of groups differ by estimation method. For instance, when comparing the parametric and nonparametric results, we notice that for blacks, the parametric estimates are generally smaller than their nonparametric counterparts. If functional form is the only issue, this would suggest that the parametric estimates are downward biased. Although the magnitude of the median partial effect is also higher in the nonparametric model for whites, this difference is smaller and hence the gap between the two groups is larger when employing nonparametric regression.

Perhaps more important than differences at the median is our uncovering substantial heterogeneity in the rates of return to schooling within groups. As stated previously, many papers allow for heterogeneity across groups, but few allow for heterogeneity within specific groups. We show substantial variation in the rate of return within groups. However, this variation is not constant across groups and/or time.

We find one striking unexpected result when examining heterogeneity in the rates of return. Specifically, we uncover significant percentages of very low and sometimes even *negative* returns for specific groups. We attempt to determine which groups have relatively low and relatively high returns by examining the characteristics that are common amongst each group. For example, we find that younger workers with young children are less likely to enjoy larger benefits of education whereas older workers with young children are more likely to have large rates of return to education. Examples of workers in the low group include both older workers and immigrants.

3 Empirical methodology

In regression we are typically concerned with predicting the left-hand-side variable given specific values of one or more right-hand-side variables. For a particular observation, this is the conditional expectation $E(y_i|x_i = x)$. The general regression model with an additive mean zero random error is written as

$$y_i = E(y_i|x_i) + u_i, \quad i = 1, 2, \dots, n.$$

Most parametric analyses assume that $E(y_i|x_i)$ is linear in x , i.e. $E(y_i|x_i) = \alpha + \beta x_i$. If this model is true and the other Gauss-Markov assumptions hold, then the estimators of α and β are the best linear unbiased estimators and one can proceed with inference and policy suggestions. However, if the true model is nonlinear and one ignores this, estimation may not only lead to inconsistent estimates, but it can also mask important heterogeneity in the marginal effects. For example, suppose the true model is quadratic in x , but one fits a linear model. In a linear model the estimated partial effect $\partial y/\partial x = \beta$ is constant for all x . Thus, not only will the linear model's result be inconsistent, but it is also ignorant of the fact that the true partial effect varies with x . Even worse, the marginal effect could take both positive and negative values. Implementing a policy based on results from the linear model when the true technology is quadratic could lead to unintended consequences for a particular group, for example, a detrimental instead of positive impact of a treatment for a sub-group of the population.

Given that the true data generating process is generally unknown, there are a few options: (1) Simply hope that the true model is linear. Given that this is only one possibility out of an infinite number of possibilities, this may be a bit naive. (2) Fit higher order polynomials as well as use interaction terms. This is a promising approach, but given the number of possibilities, it is difficult to model all of these without quickly running out of degrees of freedom. Other issues with this approach in this setting will be discussed later. (3) Let the data tell the form of the technology. This is the approach taken in this paper.

3.1 Ordinary least squares

Although OLS is well understood by economists, we feel the need to briefly describe the estimator in this particular framework. The typical Mincer (1974) regression model for

males is given by

$$\ln y_i = \alpha + \beta s_i + \gamma t_i + \delta t_i^2 + u_i, \quad i = 1, \dots, n \quad (3.1)$$

where y is annual earnings, s is the number of years of education and t represents the years of experience.³ α , β , γ and δ are parameters to be estimated and u is the standard additive error term.⁴ Note that even though our left-hand-side variable is measured in logs and experience enters in quadratic form, this model is linear in its parameters and thus may be estimated by OLS.

Of particular interest in this paper is the coefficient attached to the schooling variable, β . It represents the partial change in $\ln y$ when s is changed by one unit. It is roughly interpreted as the percentage change in earnings when schooling is increased by one year. This value is fixed for all levels of schooling which means that it is assumed that a one year increase in schooling brings about the same percentage change in earnings regardless of the number of years of schooling. Further, the model assumes that schooling is linearly separable from the other regressor(s) in the model. In other words, this coefficient is assumed constant across groups/individuals. This model is subject to the same criticism as other parametric models. Specifically, misspecification of the conditional mean will generally lead to inconsistent parameter estimates and potentially inappropriate policy prescriptions. The literature sometimes allows for more flexible earnings function by adding higher-order polynomial terms and interactions. This practice, however, does not necessarily provide a good approximation to the underlying relation between earnings and education.⁵ More important, they do not allow for within-group variations in returns to education. For example, Card (1999) points out that even a high-order polynomial parameterization of the Mincer model does not fit the age profiles well for different educational groups, and that models allowing for more flexible interactions between education and experience are needed. This view is also shared and supported by Heckman, Lochner, and Todd (2003). The authors find that “assumptions of linearity in schooling and separability between schooling and experience” are

³ Throughout the paper we will treat the right-hand-side variables as exogenous. We discuss this choice in more detail in Section 7.

⁴ This is regression P(1) of Table 5.1 (p. 92) in Mincer (1974). There Mincer gives an interpretation of the coefficients (p. 91) but it is important to note that he provides other possible functional forms because he realized first that rates of return can vary across the population and second that linearly declining post-school investment, as is assumed in P(1), need not hold.

⁵ Based on Becker type of optimal schooling models, Card (1995, 2001) show that even the simplest linear specification for marginal return to schooling implies a quadratic relation between observed education and earnings outcomes. As noted in Card (2008) and Rau Binder (2006), relaxation of the linearity assumption leads to a even more general earnings function, which could be potentially highly nonlinear depending on the underlying functions for marginal cost and return.

far more important sources of misspecification in the Mincer model than higher order terms in experience on which the literature tends to focus. Again, although we are aware that more sophisticated versions of (3.1) are available in the literature (Heckman and Polachek, 1974; Heckman, Lochner and Todd, 2008), this form is still the standard when analyzing the rate of return to education.

3.2 Generalized kernel estimation

The basic idea behind nonparametric regression is to estimate the unknown conditional mean. Here we consider a variant of the local-linear least-squares (LLLS) estimator (Fan and Gijbels, 1992; Pagan and Ullah, 1999).⁶ Specifically, we use Generalized Kernel Estimation (Li and Racine, 2004; Racine and Li, 2004) to estimate the conditional mean and gradient. To begin, first consider the nonparametric regression model

$$y_i = m(x_i) + \varepsilon_i, \quad i = 1, \dots, n \quad (3.2)$$

where y_i is the left-hand-side variable measured for observation i . $m(\cdot)$ is the unknown smooth function with argument $x_i = [x_i^c, x_i^u]$, x_i^c is a vector of continuous regressors, x_i^u is a vector of regressors that assume unordered discrete values, ε is an additive error, and n is the number of observations. In our application, y is log annual earnings and x^c contains $q = 2$ elements: years of education and experience. x^u contains a single element for whether or not the individual was top-coded.

Taking a first-order Taylor expansion of (3.2) with respect to x yields

$$y_i \approx m(x) + (x_i^c - x^c)\beta(x) + \varepsilon_i \quad (3.3)$$

where $\beta(x)$ is defined as the partial derivative of $m(x)$ with respect to x^c . The LLLS estimator of $\delta(x) \equiv (m(x), \beta(x))'$ is given by

$$\widehat{\delta}(x) = (X'K(x)X)^{-1}X'K(x)y, \quad (3.4)$$

where X is a $n \times (q + 1)$ matrix with i th row being $(1, (x_i^c - x^c))$ and $K(x)$ is a diago-

⁶ In short, LLLS performs weighted least-squares regressions around a point x with weights determined by a kernel function and bandwidth vector. Specifically, more weight is given to observations in the neighborhood of x . This is performed over the range of x and then the unknown function is estimated by connecting the point estimates. Some of the benefits of LLLS are that it requires no assumptions on the underlying functional form and allows for heterogeneity in the partial effects. Further, if indeed the true functional form is linear, the LLLS estimator nests the OLS estimator when the bandwidth is very large.

nal n matrix of kernel weighting functions for mixed continuous and categorical data with bandwidth parameter vector h (Li and Racine, 2007).⁷ Closer inspection of the estimator in (3.4) shows that the estimate is specific to x . In other words, we obtain a fitted value and derivative estimate (for each regressor) for each x . This allows us to observe heterogeneity in the partial effect of schooling.

3.2.1 Bandwidth selection

Estimation of the bandwidths (h) is typically the most salient factor when performing nonparametric regression. For example, choosing a very small bandwidth means that there may not be enough points for smoothing and thus we may get an undersmoothed estimate (low bias, high variance). On the other hand, choosing a very large bandwidth, we may include too many points and thus get an oversmoothed estimate (high bias, low variance). This trade-off is a well-known dilemma in applied nonparametric econometrics and thus we usually resort to automatic determination procedures to estimate the bandwidths. Although there exist many selection methods, one increasingly popular method is Hurvich, Simonoff and Tsai's (1998) AIC_c criterion. This method chooses smoothing parameters using an improved (in terms of bias) Akaike Information Criterion. One benefit of this method is that it tends to avoid undersmoothing which often happens with other popular methods such as least-squares cross-validation.

3.2.2 Estimation of the density of the partial effects

A benefit of nonparametric kernel methods is that they give a plethora of results. Observation specific estimates can be obtained for each regressor in a local-linear regression, implying we have $n \times q$ partial effects. It is often difficult and/or impractical to present this many values in a paper. Therefore researchers often devise ways to present the results. Some authors simply look at the mean or median of the estimates for a particular regressor. However, this ignores possible heterogeneity in the estimates. One increasingly popular method to present the results is to plot kernel densities of the estimates. This allows us to examine the entire

⁷ The generalized product kernel function for a vector of unordered and continuous variables is the product of a kernel function for continuous variable(s) and a kernel function for unordered variable(s). For unordered variables, the kernel function utilized is the one proposed in Aitchison and Aitken (1976); for continuous variables, the kernel function is second order Gaussian kernel. See Li and Racine (2007) for more details.

set of estimates for a particular regressor in one simple-to-view figure.⁸ We analyze both methods.

4 Data

The data are obtained from the Integrated Public Use Microdata Series (IPUMS) 1940 - 2000 and the American Community Survey (ACS) 2005. IPUMS data 1940 - 2000 are based on U.S. Decennial Censuses.⁹ For 1940 - 1970, the 1% samples (1-in-100 national random sample of the population) are available; for 1980 - 2000, the 5% samples (1-in-20 national random sample of the population) are available. Since Census data are conducted every ten years, we also pool together ACS 2005 (1-in-100 national random sample of the population) for our analysis to reflect recent trends.¹⁰ Our sample is restricted to male workers aged 16 and above. We focus on individuals from “regular households” and “additional households under the 2000 definition”.¹¹ Individuals living in group quarters are excluded from the sample. Moreover, we keep only the observations with unaltered or logically edited values. The observations that are manually edited using hotdeck, colddeck, or the unspecified allocation method are excluded from our analysis.¹²

To perform the analysis, three variables are of primary interest: individual earnings, years

⁸ Plotting kernel densities of predicted values and/or derivatives is analogous to that for a simple vector of data. Let $\hat{\beta}_i = \hat{\beta}(x_i)$, then the kernel density estimate for the estimated derivative is defined as

$$\hat{f}(\hat{\beta}) = (nh)^{-1} \sum_{i=1}^n K\left(\frac{\hat{\beta}_i - \hat{\beta}}{h}\right), \quad (3.5)$$

where h is a scalar bandwidth and $K(\cdot)$ is the kernel function. Stated loosely, a kernel density estimate can be thought of as a smoothed histogram. The kernel function determines the shape of the bumps and the bandwidth controls the degree of smoothness. Throughout the paper, when presenting kernel density estimates, we employ a Epanechnikov kernel and use 2.5 multiplied by the adaptive rule of spread (Silverman, 1986, pp. 47).

⁹ All questions were asked for the previous calendar year. For instance, reported earnings in the 1980 census are earnings in 1979.

¹⁰ All the data files can be downloaded on <http://usa.ipums.org/>.

¹¹ Prior to 2000, households that contain 10 or more individuals unrelated to the household head were classified as group quarters, instead of regular households. The definition of group quarters has since changed. For the 2000 census and all ACS and PRCS samples, “housing units were classified as group quarters only if they belonged to a list of such units.” (<http://usa.ipums.org/usa-action/variableDescription.do?mnemonic=GQ>) Therefore, IPUMS coded this group of individuals under group quarters and define it as “Additional households under the 2000” definition. We wanted to stick to the original Census’s definition and thus added this group back to the household.

¹² There are three ways that IPUMS deals with missing/inconsistent values. Most variables in the IPUMS have been edited for missing, illegible or inconsistent values. It is not obvious that editing will necessarily add more accurate information.

of schooling and potential experience ($age - schooling - 6$).¹³ The definition of individual total income (the sum of all sources of earnings) is different across years, as more detailed income categories are asked in later years. The only income category that is comparably defined across all years in our analysis is wage and salary income. Ideally, to separate income effects from labor supply effects, hourly wage is needed. Information on weeks and hours worked is, however, not consistently available. The variable “weeks worked in the previous year” is continuous in all years except in 1960 and 1970 where it is documented in intervals. The variable “hours usually worked per week in the previous year” is not available before 1980. This is the same problem Mincer and others faced and therefore, we construct the measure of individual income based on individual’s annual wage and salary income, as did Mincer.¹⁴ To ensure comparability across years, we also adjust the wage for inflation using the adjustment factors provided by IPUMS.¹⁵ Individuals with negative income are excluded. Moreover, the wage variable is top-coded with different topcodes across years. Therefore, we create a dummy variable, equal to one if income exceeds the topcodes and zero otherwise, that is included in the analysis whenever appropriate.¹⁶

Another issue is that schooling is not measured in a consistent fashion over time. Prior to 1990, the census asked individuals how many grades of school or years of college they had. Starting from 1990, however, the census and ACS asked individuals about the highest grade or diploma completed. Moreover, grades completed below 9th grade are reported in three-year intervals. Following the literature (e.g. Lemieux and Card, 2001), for data from 1990-2005, we replace these intervals with midpoints.¹⁷ Our definition of potential working experience equals current age minus years of schooling minus 6.¹⁸ Negative values are recoded as zero.

After the initial estimation, a number of other variables are used to further investigate potential heterogeneity in the return to education both across and within groups. These

¹³ Actual experience is not available in the data required for this study. Further we concentrate only on males for which potential experience mimics actual experience very well. In addition, the preponderance of current rate of return studies use potential experience which enables comparison between our results and others’.

¹⁴ Any values exceeding 999999 are recoded as missing values.

¹⁵ See IPUMS website for consumer price index adjustment factor for each year <http://usa.ipums.org/usa-action/variableDescription.do?mnemonic=INCTOT>.

¹⁶ See IPUMS website for topcodes for each year <http://usa.ipums.org/usa-action/codes.do?mnemonic=INCWAGE>.

¹⁷ The coding scheme is as follows. No schooling completed (0); Nursery school (0); Kindergarten (0); 1st to 4th grade (2.5); 5th grade to 8th grade (6.5); 9th grade (9); 10th grade (10); 11th grade (11); 12th grade, no diploma (12); High school graduate (12); Some college (13); Associate degree (14); Bachelor’s degree (16); Master’s degree (18); Professional degree (20); Doctorate degree (20).

¹⁸ It is also possible to treat age and experience as discrete variables in our estimation procedures. However, to anchor our results to previous studies, we treat these variables as continuous.

variables include a dummy variable indicating whether or not the worker is an immigrant, a dummy variable indicating self-employment status, region dummy variables (Midwest, South, West, and East), a dummy variable indicating whether or not there are any children younger than five in the household, a dummy variable indicating if an individual is married, and dummy variables indicating whether an individual is white or if an individual is black.

4.1 Sample size

Nonparametric estimates are desirable because they are able to obtain a separate estimate for each unique set of realizations of the regressors. One down side to this type of estimation is the computing time. Estimating the bandwidth vector is extremely computationally expensive. Further, we plan to consider many regressions for many years. In order to minimize the computing time we chose to only sample 25,000 observations for any particular regression.¹⁹ Thus, the samples (likely) do not include the same individuals across regressions, but we have the same individuals when comparing across estimation techniques (parametric vs. nonparametric). We also ran the results for some models with much larger data sets and did not find significantly different qualitative results. These are available upon request.

4.2 Sensitivity to Labor Supply Effects

We focus on annual wage and salary income since information on weeks and hours worked is not consistently available across years. Because individuals with more education may also work more, we re-estimate our models using hourly wages for those years when this variable is available. The main conclusions are qualitatively unchanged, and are therefore omitted but are available upon request. In the interest of brevity, we simply highlight the main differences between the two potential left-hand-side variables before going on in the next section to describe the annual wage and salary results in detail. First, both median estimates and the dispersion of rates of return based on annual wage are in general larger than those based on hourly wage, confirming our hypothesis that education and hours worked may be positively correlated. Second, we also find that in 1990, individuals aged 30-39 had the highest rate of return. Moreover, while individuals aged 30-39 had the smallest dispersion of the return to education in 1940, 1980, and 1990, individuals aged 20-29 had the smallest dispersion in 2000 and 2005. Finally, third, we also find that the convergence in the rates of return

¹⁹There are a few cases when IPUMS did not have 25,000 observations for a particular group. Specifically, these were for the sample of Black (1940-1970), Single (1950), Immigrant (1950-1970), 20-29 (1950), 30-39 (1950), 40-49 (1950) and 50-59 (1950). In these cases we took the entire sample available in IPUMS.

between immigrants and natives continued, and that eventually the return to education for immigrants exceeded that for natives by 0.6 percentage points in 2005. Otherwise, the results are qualitatively similar.

5 Results

Table 1 gives the baseline parametric and nonparametric results over time both for the pooled sample and for specific groups. All regression estimates in bold are those that are significant at (at least) the 10% level. Tables 2 and 3 give the breakdown by cohort for the same groups in the previous table. Finally, Table 4 gives schooling rate of return estimates by education level. We report the median of the nonparametric returns in all the tables, as they are the most comparable to the parametric results.²⁰ However, given that our nonparametric estimates are specific to a given value of x , we present the coefficients of variation of the returns as well in the lower panel (Panel B.2). Figures 1-5 complement these tables by displaying the density plots of the estimates for both the pooled sample and specific groups over time.

5.1 Pooled results

We first present the pooled parametric results separately for 1940 to 2005 in the first column of Table 1, Panel A. Regression estimates imply an undoubtedly positive return to education in the labor market, ranging from 8% to 13%. The results reflect the long-run trend in rates of return to education. Comparing the estimates from 1950 with those from 2005, the results indicate an increase in the return to education of 4 percentage points in the past fifty years. The increase was, however, not monotonic over time; the return to an additional year of schooling increased from 8.4% in 1950 to 9.1% in 1960, then fell to 8.3 in 1970 and further to 8% in 1980. Afterwards, the rate of return started to increase in 1990, stagnated in 2000, and continued to increase in 2005.

Turning to the nonparametric estimates (the first column of Table 1, Panel B.1), we first notice that each of the median estimates is large and positive. Unlike their parametric counterparts, nonparametric results show a clear monotonically increasing trend since 1950; the estimate was 8.2% in 1950 and then increased to 9.3% in 1960. Afterwards, the number

²⁰We give the median value of the estimated partial derivative of the conditional mean with respect to schooling of the nonparametric estimates. The median is a more robust measure as the mean is sensitive to outliers. Further, we present the entire distributions in Figures 1-5.

remained relatively constant, from 9.3% in 1960, to 9.4% in 1970, and to 9.6% in 1980. The rate of return started to grow more quickly in 1990; the return was 13.3% in 1990 and rose to 14.3% in 2005.

Both parametric and nonparametric estimates indicate lack of significant growth during the period of 1960 - 1980. To those unfamiliar with the literature, this can be partly explained by at least two phenomenon during the same period. First, there was a sharp increase in the supply of educated workers during that time period (Angrist and Chen, 2007). Second, the recessions in the 1970's and early 1980's likely decreased the demand for labor.

While these trends are more or less consistent with the literature (Goldin and Katz, 1999; Card and Lemieux, 2001), the nonparametric estimates of the returns to education across years indicate that there are two important features of the trends that are masked by the parametric OLS estimation. First, comparing the parametric and nonparametric results (assuming functional form is the only issue) indicates that the linearity assumption imposed in the parametric estimation leads to underestimated rates of returns across years, except in 1950. For example, the parametric estimate of the rate of return is 10.8% in 2000 while its nonparametric counterpart is 13.5% – a difference of 2.7 percentage points. Moreover, the discrepancy between parametric and nonparametric results also imply that, although the growth of the rate of return to education slowed down in the 70's and early 80's, the overall negative impact on the labor market (sharp increase in the supply of educated workers and the recession during the same period) may be exaggerated due to potentially biased parametric estimates.

Second, although lack of significant growth during the period of 1960 - 1980 suggests that worsened economic conditions during the same period may contribute to this phenomenon, there is, however, no reason to believe that a recession will affect younger and older, or black and white, or native and immigrant workers equally. Imposing constant returns to education as in OLS estimation ignores such a problem. Partly for this reason, we examine potential heterogeneity in the rate of return to education. Recall that the nonparametric estimation method allows us to obtain observation-specific returns to education for each individual. We report two sets of results regarding the distribution of returns to education. First, Table 1, Panel B.2 shows the dispersion of the returns to education. Not only does this measure indicate the existence of heterogeneity in returns to education within each year; it also provides an important indicator of the extent of underlying potential uncertainty for educational investment. We can see that the dispersion of the returns to education was steadily declining, from 0.872 in 1950 to 0.373 in 2005. This result confirms the fact that

changes in economic conditions impact individuals differently. Second, given the observation-specific estimates obtained, Figure 1 plots the densities of the nonparametric estimates. None of these densities appear to be normal and there appears to be a significant amount of skewness.²¹ Specifically, there are many individuals whose rates of return lie to the right of the mode(s). Another interesting observation is that while the dispersion of the distribution has decreased over time, there are still two large extremes across years. First, a small fraction of estimates lie below zero. These density plots imply that while the majority of the population certainly benefits from education, some individuals get no benefit at all or even have negative rates of return. Second, some individuals receive much larger rates of return than others. Later we will investigate the characteristics of individuals receiving extremely low (bottom 5%) and high (top 5%) rates of return to education in more detail.

5.2 Sample splits

As discussed above, it is naive to believe that all groups of individuals receive the same rate of return to education. One way for parametric estimation to allow for heterogeneity is to estimate the model for different sub-samples, separately. Thus, as opposed to taking a random sample of 25,000 workers from the population, we take 25,000 observations for those belonging to each specific group of interest. Specifically, we consider four age groups, two racial groups, as well as splits (stratifications) by immigration status.²² Again, for each of these ten groups we sample 25,000 observations each year. We then run OLS and nonparametric regressions on each (cross-sectional) group separately for each year and report the results in Table 1 (columns 2-9).

5.2.1 Age groups

The second through fifth columns of Table 1, Panel A present the parametric results for four age groups (20-29, 30-39, 40-49 and 50-59). We again find a large, positive return to education, that is both economically and statistically significant. The estimates vary from about 6% to about 19%. In line with the pooled parametric results above, the results also show that there has been a long-run increasing trend in the return to education, although such a trend is non-monotonic. The return to education is higher for younger workers than older workers, and there is a monotonic decline in the return to education as age increases.

²¹ Formal tests for these type of phenomena can be found in Chapter 12 of Li and Racine (2007).

²² We also estimate the models separately for marital status. These results also show large heterogeneity in returns to education and thus are omitted but available from the authors upon request.

For example, in 1940, the return was as large as 18.8% for workers aged 20-29 and fell to 14.3% for workers aged 30-39, to 10.8% for workers aged 40-49 and finally to 8.1% for workers aged 50-59. Similar patterns are also found for other years. Another feature is that the gap in the return to education between younger and older workers has generally been widening (but not consistently) over time since 1950. For example, the difference in the return to education between workers aged 20-29 and workers aged 50-59 increased from about 4.5% in 1950 to roughly 10.9% in 2005.

In Table 1, Panel B.1, columns 2-5 give the median of nonparametric estimates of rates of return to education by age group. There are substantial discrepancies between the parametric and nonparametric results. First, during the period 1940 – 1960, the nonparametric estimates are in general larger than the parametric estimates, confirming the pooled results above. However, during the period 1970-2005, we find that the nonparametric estimates are consistently smaller than the parametric ones for individuals aged 20-29; and that the nonparametric estimates are in general larger than their parametric counterparts for older workers. Altogether, the results imply that the wage structure may be different for younger workers; and that the wage structure could change over time. Imposing a linear wage structure could thus lead to either downward or upward biased estimates in the rates of return to education, depending on the underlying wage structure. These highlight the importance of employing the nonparametric estimation method which is free of restrictive assumptions on the functional form.

Second, as mentioned, one point of interest is the drop in rates of return in 1980. We argued that the decline can partially be explained by both an influx of educated labor and the relatively weak economy. We consistently find that younger and older workers were not impacted equally by the worsened labor market conditions during the period 1970-1980. In particular, the most impacted group was individuals aged 20-29 (a decrease of 2.2 percentage points), followed by individuals aged 50-59 (a drop of 1.3 percentage points). Individuals aged 40-49 were hardly affected, and those aged 30-39 actually had an increase of 2.1 percentage points. The individuals aged 20-29 were also most affected by the draft (Angrist and Chen, 2007) and have the least work experience to fall back on. The large drop for older workers (50-59) is likely related with retirement trends (Soldo, Mitchell, Tfraily, and McCabe, 2006; Clark and Mitchell, 2005). Comparing these results with their parametric counterparts, we can see that the detrimental impact of worsened labor market conditions on young workers is underestimated.

In addition to *between-group* heterogeneity across age groups, we also find a large *within-*

group heterogeneity in returns to education, which cannot be revealed by the OLS parametric estimations. In particular, we find that the dispersion of the return to education is consistently larger across years for individuals aged 20-29 than that for older groups. The pattern of the dispersion across groups is, however, unclear. In particular, while individuals aged 30-39 had the smallest dispersion of the return to education in 1940, 1960, 1990, and 2005, individuals aged 50-59 had the smallest dispersion in 1950, 1970, 1980, and 2000. Figure 2 displays the density plot of these estimates. The plots show that the densities are more likely to cover both negative and positive values for younger workers (aged 20-29) than older workers.²³ In other words, additional education could be harmful to some young workers. For example, approximately 20% of workers aged 20-29 received negative returns in 1990. On the other hand, a large fraction of individuals aged 20-29 received extremely large positive returns. For example, in 1990, the maximum return for individuals older than 30 was 20%, whereas roughly 10% of individuals aged 20-29 had returns higher than 20%. Moreover, the densities for younger workers, in general, appear to be flatter and to have fewer modes.

5.2.2 Race

The parametric results for white and black workers are presented in the next two columns of Table 1, Panel A. While we continue to find a positive return to education for both groups, years of schooling are rewarded differently for white and black workers. Except in 1940, black workers consistently enjoy larger rate of return benefits than their white counterparts (Welch, 1973; Card and Krueger, 1992). Another point is noteworthy. We continue to find a long-run, although non-monotonic, increasing rate of return for white workers. The pattern for black workers is, however, not consistent with the pooled results above. In particular, while we also find a slight drop in 1970, rates of return quickly started increasing again, from 10.1% in 1980, to 13.2% in 1990, and to 15.9% in 2005.

Moreover, this increase for black workers is also larger than that for white workers. Despite the increasing trend for black workers, there existed a gap in the rate of return to education between black and white workers. The gap tended to converge before 1970, but started to diverge after 1970. For example, the difference in return to education between black and white workers was 1.4% in 1950, and slightly decreased to 1.3% in 1970, but rose to about 4% in 2005.

Turning to the nonparametric estimates, we again find that black men have higher rates of return than do white men, with the largest disparity being 10 percentage points in 2005.

²³This likely partly reflects those young people who are working, but still in school and those just out of school.

Comparing parametric and nonparametric results, several discrepancies stand out. First, for both blacks and whites, parametric estimates are generally smaller than their nonparametric counterparts, indicating that the parametric estimates are downward biased. Second, even though, similar to the parametric results by race, we again find that the time trends for both blacks and whites are not consistent with the time trend from the pooled nonparametric results – where there exists a clear monotonically increasing trend since 1950 – the implications of the differences between stratification and pooled results are completely different. On the one hand, the parametric results suggest that both black and white workers were affected during the period 1960-1970, but that whites were more impacted by the worsened labor market and economic conditions than blacks, consistent with the literature (see, e.g. Kniesner, Padilla and Polachek, 1978). On the other hand, the nonparametric estimates suggest the opposite. In particular, black workers were not at all impacted during the period 1960-70. Instead, there was a steady increase in their returns to education during this period. Furthermore, although white workers were impacted during the period, the magnitude of the impact is exaggerated by the parametric estimates. Finally, the nonparametric estimates suggest that black workers, on average, were impacted negatively in 2000 (possibly by the dot-com crash), a fact masked by the parametric estimates. In particular, the returns to education decrease from 20.5% to 19%, a decrease of 1.5 percentage points.

Nonparametric estimates allow us to obtain the distribution of returns to education for both blacks and whites. Looking at the results for the dispersion of the returns to education, we confirm that there exists a large heterogeneity in returns to education even within each racial group. Another surprising yet interesting result arises. We find that the dispersion of the return to education is smaller (larger) for blacks than for whites during the period 1960 – 1990 (1940 – 1950 and 2000 – 2005). In other words, not only were black workers less impacted by worsened economic situations, on average; they were also facing less uncertainty in the labor market in the bad economy! A complete characterization of the distribution of returns to education across racial groups is displayed using the density plots of these estimates in Figure 3. The results indicate that in most years, there existed some whites having negative rates of return. For example, in 1980, 15.2% percent white workers had negative returns. On the other hand, there was only a small fraction of blacks receiving negative returns, except in 1950 (roughly 7 percent). Another finding is worth mentioning. Since 1980, there has been a considerably large fraction of blacks receiving extremely large positive returns. For example, in 2005, 25 percent of black workers received the return of more than 30%. In the same year, about 20 percent of blacks had the maximum return

received by whites (33.2%).

5.2.3 Immigration status

The parametric results by immigration status are presented in the final two columns of Table 1, Panel A. The results indicate that the rate of return to education is considerably lower for immigrants than for natives (Chiswick, 1978; Chiswick and Miller, 2004; Rivera-Batiz, 2007). For example, in 1940, while the return to education for natives is as large as 15%, it is only 6% for immigrants. Patterns for immigrants and natives also differed. In particular, rates of return for immigrants increased monotonically. The rates of return for natives first increased in 1960, but then decreased in 1970 and 1980, although the rates of return started increasing again in 1990. Because of such different patterns between immigrants and natives, the growth in the rate of return to education for immigrants was much faster than that for natives before 1990, closing the gap to some extent over time. In particular, the difference was about 4 percentage points in 1950, but fell to 0.2 percentage points in 1980. After 1990, the growth slowed down for immigrants but rose for natives, leading to a larger gap of nearly 4 percentage points in 2005.

The final two columns of Table 1, Panel B.1 present the nonparametric results by immigration status. First, the nonparametric results indicate that the median return to education is lower for immigrants than for natives in all years, consistent with the parametric results. However, we notice that for immigrants, parametric estimates are smaller than their nonparametric counterparts. Such a discrepancy indicates that the return to education is consistently underestimated for immigrants in the parametric model.

Nonparametric results also show a different picture of the trend over time. Except in 1940, the return to education was relatively stagnant for natives during 1950 – 1980, whereas there was a steady increase for immigrants from 1950 to 1980; eventually the gap in the return to education between immigrants and natives was narrowed to .1 percentage points. However, such a trend did not persist. After 1980, the gap between immigrants and natives appeared again; the gaps were about 2 percentage points in 1990 and 2005. Second, while the return to education for both immigrants and natives fell in the 1970's and early 1980's, the timing is different; the decrease occurred in 1970 for natives and in 1980 for immigrants. Such a difference implies that the economic conditions impacted immigrants and natives differently. Third, the dispersion of the rates of return to education is in general larger for immigrants than for natives. Although education acquired abroad (i.e. out of the U.S.) is not directly transferable for all immigrants, because education quality or curriculums vary greatly across

countries, the extent of transferability may also be different across the countries of origin (Friedberg, 2000).²⁴ This implies that returns to education can differ among immigrants. On the other hand, although quality and curriculums also vary across states within the U.S., the variation of quality or curriculums for education received in the U.S. should be much smaller relative to the variation for education received abroad in general. As a result, although there exists heterogeneity in returns to education for both immigrants and natives, the observed dispersion of the returns to education is, however, larger for immigrants than for natives.

In terms of the density plots (Figure 5), two points stand out. First, the densities were more likely to cover the negative regions in early decades than in recent decades. For example, approximately 37 percent of immigrants had negative rates of return in 1950, but only 1.4 percent of immigrants had negative rates of return in 2000 and 2005. Taking a look at the second panel of Table 3 shows that the dispersion of the rates of return to immigrants dropped dramatically after 1980 and thus likely explains this result. However, the question remains as to why this dispersion fell. Second, while the dispersion is generally larger for immigrants than for natives (except in 1980 and 2000), immigrants are less likely to exhibit large returns than natives.

5.3 Cohort results

Comparison between age groups in a particular year may not necessarily reflect the life cycle pattern for returns to education. A better understanding of the life cycle pattern can be obtained by tracing the return to education for the same (synthetic) cohort over time (Card and Lemieux, 2001). The results for cohorts by group are presented in Tables 2 and 3.

5.3.1 Race

The parametric results by race are presented in the Panel A of Table 2. First, we again see that younger workers have higher returns to education than older ones, consonant with the results above. Also, comparing the results for black workers with those for white workers, we again find that education is in general more rewarding for black workers than for white workers. But we do find three exceptions where older black workers have slightly lower

²⁴For example, using a research design similar to the existing literature for other countries such as U.S., Pischke and von Wachter (2008) find zero returns to an additional year of compulsory schooling, in stark contrast to the literature. Their results suggest the contents of the curriculums differ across countries; in particular, German students may be much better prepared in basic labor market relevant skills before the end of compulsory schooling (9th grade) than students in other countries such as U.S. and U.K..

returns to education than do their white counterparts. For example, black individuals aged 50-59 received 4.9% while their white counterparts had 5.4% in 1980.

Second, looking at the return to education for the same cohort over time (diagonals), we find that the return to education for the same cohort is, in general, higher in their 20's than in their later lives. Such a result holds for both black and white men. The exceptions are mostly for those in 1980 when a decline in the rate of return to education occurred. Such exceptions are masked by comparing the age groups in the same year, and thus highlight the importance of looking at the cohort effects.

The decline in the rate of return to education over the life cycle is usually larger for blacks than for whites. While the difference between black workers in their 20's and the same cohort in their 50's is around 6%, the difference between white workers in their 20's and the same cohort in their 50's is roughly 4%.

The nonparametric results by race are presented in the Panel B of Table 2. Again, splitting the results further by age groups reveals that the advantageous situation in the labor markets for blacks mainly comes from younger blacks (aged 20-29 and 30-39). For individuals in these groups, blacks generally have higher rates of return. On the other hand, compared to the parametric results, we find more exceptions for older individuals (aged 40-49 and 50-59) where whites actually enjoy larger benefits from education than blacks.

We again find that individuals generally have larger rates of return when they are in their 20's than when they are in their 50's, consonant with the parametric results above. This result holds for both blacks and whites. However, the pattern of decline is different for blacks and whites. The decline in rates of return over the life cycle appeared to be monotonic for blacks, although there existed two exceptions in 1970 and 1980. The exceptions could be explained by the fact that individuals were impacted by the worsened labor market and economic conditions during the period. For whites, the pattern is non-monotonic. For white workers aged 20-29 in 1970 and 1980, the monotonic decline over the life cycle did not start until their 30's; since they were more likely impacted by the worsened economic conditions at the time, rates of return were relatively low to begin with. For white workers aged 20-29 in 1940 and 1950, rates of return appeared to decline when they were in their 30's, but increased in their 40's, and eventually declined again in their 50's. Moreover, comparing the rates of return in individuals' 20's and 50's, we again find that the rates of return in general decline faster for blacks than for whites.

Comparing the cross-sectional and cohort-based estimates, we find that the cross-sectional estimates in general overestimate the decrease in the return to education over the life cycle

for both whites and blacks.

5.3.2 Immigration status

The parametric results by immigration status are presented in the Panel A of Table 3. First, both the cross-sectional (Table 1) and cohort-based (Table 3) estimates indicate a decrease in the return to education over the life cycle. However, the cross-sectional approach seems to overestimate (underestimate) the magnitudes of the decline for immigrants (natives). Second, consistent with the results above, the estimates imply that the rate of return to education is in general lower for immigrants than for natives. Third, the results indicate that the life cycle decline in the rate of return to education is smaller for immigrants than for natives.

The nonparametric results by immigration status are presented in the Panel B of Table 3. We again find that natives in general have a larger return to education than immigrants. However, the implied patterns over the life cycle are different for natives than for immigrants, which is masked by the parametric results. For natives, we generally find a declining return to education over the life cycle, although the pattern is not necessarily monotonic. By contrast, for immigrants, the pattern is in general reversed. Not only are there quite a few cases where individuals have lower rates of returns in their 20's than those in their 50's; there also appears to be some cases where the rates of return increased over the life cycle and peaked in their 40's. This is an interesting result. There are two possible explanations. First, it may be explained by changing composition of immigrants over time. For example, immigrants who do not perform well in the labor market (e.g. have low returns to education) may return to their home countries, because of the bleak earnings prospect in the U.S. As a result, the increase in the returns to education over time may simply reflect the fact that individuals with better economic gains stay in the U.S. Second, this result could also be explained by immigrants' assimilation process. Upon arrival in the U.S., immigrants have lower returns to education because their education is not easily transferable; assimilation to the local society – either by learning the culture or by acquiring U.S. education – can increase the return to home-country education among immigrants. For example, Friedberg (2000) finds that “acquiring additional education following immigration appears to confer a compound benefit by raising the return to education acquired abroad.”

5.4 Results By Education Level

A question of interest is whether there exists diminishing returns to education, which implies that the rate of return is a monotonically decreasing function of education level. In order to investigate this question, we run the same parametric regressions augmented by interaction terms between years of schooling and a set of dummy variables (High school and below, High school, some college, and college and above) separately for each year. The results are presented in Table 4. The patterns varied widely across years; none of the results was completely consistent with the pattern implied by diminishing returns to education, although we do find that rates of return for individuals with education below high school are consistently larger than those with college education and above. However, the restrictive linear form imposed in parametric estimation could be a possible reason why we do not see a clear pattern here.

Turning to nonparametric results, we again consistently find that individuals with college education and above have lower rates of return than those with education below high school. These results are more or less consistent with the pattern implied by the hypothesis of diminishing return to education. Comparing the parametric and nonparametric results, rates of return for individuals with college education and above are consistently overestimated in the parametric estimation when imposing the linear form in parametric estimation, impeding us from seeing the underlying patterns.

One may suspect that the pattern found here may be attributed to the labor supply effects of education level. Thus, we also conduct the same analysis by years of school when information on hourly wages is available. The results are not sensitive to adjusting for labor supply effects, although the adjusted estimates are in general smaller than the estimates without adjustment, as expected. The results are thus omitted here but available from authors upon request.

We also uncover a large heterogeneity in returns to education within each education level. There are two possible explanations. First, it could be due to “sheepskin effects”. Given the same years of schooling, some individuals may have different degrees (see, e.g., Park, 1994, Table 3). Second, this could be better understood using Becker type of optimal schooling model: an individual’s optimal schooling is determined by their potential returns to education (benefits) as well as their marginal costs. Two individuals with different marginal returns to education could choose the same level of education because they face different educational costs. As a result, even for individuals with the same level of education, returns to schooling could be different.

6 Extreme returns

The final set of results are found in Tables 5 and 6 which look at the characteristics of individuals who possess extremely large and small returns in order to figure out which characteristics of individuals are more likely to be associated with extreme returns to education. As we have seen above, there exists considerable heterogeneity both across groups and within groups. Some individuals enjoy much greater rewards to education than others. Some individuals are disadvantaged in the labor market, receiving extremely low returns to education. Even more dramatic is the discovery of negative rates of return to education for *some* individuals.²⁵ In this sub-section we attempt to uncover who possess extreme values of rates of return and what characteristics they have in common.

6.1 Who is Among the Top 5%?

The lengthy Table 5 tries to uncover whether particular groups have extremely large rates of return to education (the top 5% of the distribution of rates of return) with respect to other characteristics. Each element in the table represents the percentage difference for the top 5% of people in the column group who have the row characteristic from the bottom 95% of people in the column group who have the row characteristic. Specifically, we take the average of the characteristic (the percentage of people in that group who have that characteristic) for the people in the top 5% and subtract the average of the characteristic for the people in the bottom 95% and test the significance (t-test) of that difference. For example, the first element in the table says that in 1940, for people aged 20-29, the difference in the percent married between those whose rates of return are extremely large and the percent married for those whose rates of return are extremely low is 15.8 percentage points. In other words, for those people aged 20-29, marrieds are more likely to have a larger rate of return to education. Further, this difference is significant at (at least) the 10% level given that the value is in bold.

The pattern for age groups is generally consistent over time. In particular, we find that the youngest individuals (aged 20-29) are more likely to have much larger rates of return (positive values in Table 5); and that older individuals are less likely to have extremely large

²⁵ Although there is much anecdotal evidence (for example, Ph.D.'s driving taxi's), to the best of our knowledge, we believe that no one has discovered negative rates of return to education for at least some individuals within the population. That being said, an occasional negative rate of return to education with nonparametric methods should not be alarming or out of the ordinary given the small sample biases of nonparametric estimators. However, if we find a large proportion of the estimates to be negative, this would suggest that these negative rates do exist.

return to education (negative values in Table 5). This result is expected given the results for the cohorts.

There seems to be no consistent pattern for the association of having a young child with the probability of receiving extremely large rates of return across groups. But there are several interesting results noteworthy for certain groups over time. First, individuals in their 30's and 40's that have a child under the age of five are less associated with a large rate of return, although there are a few exceptions (1960 and 1970 for individuals aged 30-39 and 1940, 1950, and 1980 for those aged 40-49). On the other hand, the pattern for the youngest (aged 20-29) and the oldest (aged 50-59) cohorts is clearer. We generally find that younger individuals having a child under age 5 are less associated with larger benefits of education, whereas older workers who have a young child are more likely to exhibit a large rates of return to education.

Second, the difference for immigrants was positive during the period 1940-1980 and turned negative after 1980, implying that having a young child makes it less likely immigrants exhibit large rates of return in recent years.

Consistent with the results above, blacks in general more readily receive much larger rates of return, although there appears to be some exceptions, most of which happened in 1970 and 1980. Another interesting finding is worth mentioning. Unlike their native counterparts, black immigrants enjoy large rates of return to a lower degree, putting them in a disadvantaged situation in the labor market.

Being an immigrant is not associated with a lower probability of receiving a large rate of return for groups such as individuals aged 20-49 during 1940-1980. Such a results is not necessarily conflicting with the results above where the median returns for immigrants are in general smaller than natives. This perhaps reflects the large dispersion of rates of return received by immigrants.

Finally, we have an interest in the region of origin for the immigrants and how this affect the probability of having extremely large rates of return. We considered several regions, but report results for the America's, Europe and Oceania, and Asia. We consistently find that Asian immigrants more readily have extremely large rates of return, and that this difference was generally statistically significant. Before 1970, if the immigrant came from the America's, then they were more likely to have a large rate of return. However, the difference has gradually decreased over time and eventually turned negative in 1970. After 1970, an immigrant from the America's became less likely to have larger returns. On the other hand, immigrants from Europe and Oceania were less likely to have larger returns before 1990

and had become more likely to enjoy greater returns since then. This difference was again generally significant. The pattern may reflect the changing composition of immigrants from different continents over time.

6.2 Who is Among the Bottom 5%?

The counterpart Table 6 tries to uncover which groups have extremely low rates of return to education with respect to other characteristics. We consistently find that older individuals (aged 50-59) are more likely to have extremely low rates of return (negative values in Table 6), and that individuals aged 30-39 are less likely to possess extremely low rates of return (positive values in Table 6). For other cohorts (aged 20-29 and 40-49), the result is less clear. During the period 1950-1970, individuals aged 20-29 are in general more likely to have extremely low rates of return, but starting in 1990, individuals aged 20-29 became less likely to be among the lower tail of the distribution of rates of return. Conversely, in the early years (1950-1990, except 1980), individuals aged 40-49 are less likely to receive extremely low rates of return, but during the period of 2000-2005, the probability of obtaining low rates of return increased for individuals aged 40-49.

For individuals aged 30-59, having a child under age 5 made it more likely to exhibit extremely low rates of return. On the other hand, having a child seems to be beneficial for other groups, for they are in general less likely to possess extremely low rates of return, except for individuals aged 20-29 in 1940 and whites in 2005. In contrast to the results for the top 5%, the results for blacks do not show a clear pattern, although we consistently find black immigrants are less likely to receive extremely low rates of return. For other groups, the results are mixed, especially during the period 1940-1970.

We find that immigrants are generally more likely to receive extremely low rates of return. Although there appears to be some exceptions for individuals aged 20-29, those aged 50-59, and singles (negative values in Table 6), most of the exceptions are not statistically significant. Finally, we do not find a consistent association between the region of origin for the immigrants and the probability of having extremely low rates of return for immigrants from Europe and Oceania and Asia. On the other hand, immigrants from the America's are in general more likely to possess low rates of return, except in 1940 and 2005.

7 Where do we go from here?

The point of this exercise was to showcase heterogeneity in the rate of return to education and to attempt to uncover who had a higher rate of return. Also, during this exercise we found that a significant proportion of the rates of return were very small or even negative. We further attempted to figure out which groups were most represented among those obtaining negative rates of return as well as who garnered the highest rates of return.

Further research should focus on allowing for the possibility of these heterogenous effects and uncovering why particular groups differ in terms of their rates of return to education. Also, we think more emphasis should be placed on heterogeneity within groups as opposed to solely across groups. In a related paper by Card and Krueger (1992), they examine how returns to education vary across states and cohorts and demonstrate how differences in school quality across states could help explain the observed heterogeneity in the returns to education. In particular, they proceed in two steps like ours. They first estimate very general models of earnings function (including the interactions of various geographical variables such as state-of-birth and region-of-residence with education) to obtain heterogeneous rates of returns to education and then link various measures of school quality to explain these estimates. Even though the models are relatively flexible, one of the disadvantages of their method is that they eliminate within-group heterogeneity in returns to schooling, which we show to be quite large in our analysis. Thus, using our estimates of returns to education (taking into account within-group heterogeneity) could shed further light on the relation between heterogeneous returns and potential causes such as school quality. It has been argued in this paper that recessions caused a drop in the rate of return to education in 1980. This suggests that labor demand related variables as well as region/state effects may also be useful in uncovering the true rates of return to education. Finally, research should also be conducted to see how these results differ in other developed as well as lesser developed and emerging countries.

One of our main findings is that individuals from disadvantaged groups tend to have larger returns to education, on average. This finding itself is not novel (see, e.g. Brand and Xie, 2010; Card and Krueger, 1992; Welch, 1973). Individuals from disadvantaged groups may be more likely financially constrained, and the large returns to education may reflect the large educational costs (Card, 2001). In addition, Brand and Xie (2010) hypothesize that *negative selection mechanism*, as opposed to *positive selection mechanism* in economics,

for education may explain this result.²⁶ Individuals from advantaged backgrounds are more likely to receive higher education, and for this group, higher level of schooling is “a culturally expected outcome”. Thus, compared to individuals from disadvantaged groups, those from advantaged backgrounds are more likely to receive *smaller* returns to education, on average. We, however, also find a large heterogeneity in returns to education even *within* these advantaged groups. This result has two implications for future research. First, is the reward structure for schooling better for black workers than for white workers? Looking at the difference in *average* returns between racial groups, we would certainly think so. However, our finding of a large within-group heterogeneity suggests that this may not be the case. Comparisons of the returns to education between blacks and whites at the distributional level perhaps using stochastic dominance tests may thus be useful in this regard. Second, the within-group heterogeneity in returns to education suggests that more than one selection mechanisms may be at work. Further investigations of the cause(s) of the within-group heterogeneity are thus needed and would help us better understand educational decisions and evaluate different theoretical models.

Finally, we want to make some remarks regarding exogeneity. Recent work examining the rates of return to education often considers potential ability biases and obtains mean rate of return estimates using instrument variable (IV) estimators. Given the strong evidence of heterogeneity we find in this paper, it is hard to argue that our results will go away once one corrects for the ability bias. For example, Harmon et al. (2003) argue that “ability bias and measurement error more or less cancel out so there is no advantage to IV”. And the ability bias does not necessarily impact the whole distribution of rates of return the same way as it may impact the mean return. Card (2008) also summarizes the literature noting that existing evidence from IV estimations are not too far away from those obtained using the simple Mincer approach. In fact, our results present a great challenge to the literature using the instrument variable approach. The development of econometric literature in instrumental variable estimation has shown that, in the presence of heterogeneity, one can only identify the average effects for the compliers, the subpopulation whose treatment is impacted by the instrument variable used (Imbens and Angrist, 1994); in the current context, the treatment is years of schooling. Therefore, the estimates obtained using as IVs school reforms or other similar natural experiments are the average rates of return for those who are most likely impacted by these instruments (Card, 2001). These estimates could be dramatically different

²⁶ Positive selection mechanism in Brand and Xie (2010) refers to the mechanism that individuals receiving larger returns to education would receive more education.

from those for the rest of the population. For example, the literature generally finds that the IV estimates are larger than their OLS counterparts. One of the reasons is that many of these reforms are enacted to lower the educational costs, and thus the individuals impacted by the reforms are likely those disadvantaged individuals with relatively high marginal cost as well as high marginal returns to schooling (Card,1995; Card, 2001). Our results confirm that disadvantaged groups such as black workers, on average, tend to have larger returns to education, providing further supporting evidence for the explanation above. However, if the compliers consist of only a small portion of the entire population, it is not quite clear why the estimates are of interest at all. To push this issue to the extreme, if one is interested in obtaining the complete picture of the underlying rates of return, one may soon run into a very difficult hurdle. Specifically, it can be argued that an instrument is needed for every subpopulation. Therefore, combining our methods with an IV approach sounds fruitful, but even if we are able to obtain an IV for each subpopulation, we still need to pay attention to potential heterogeneity within each subpopulation.

Another interesting, related question arises: Who are those individuals who are likely to be impacted and have large returns to education? That is, what characteristics do they possess? Employing the procedure laid out in Section 6, along with survey data with more detailed information on personal and family characteristics may shed light on these questions. Answers to these questions could help design more effective policies aimed at increasing individuals' education and improving their labor market performances.

8 Conclusion

In this paper we examined heterogeneity in rates of return to education both across and within groups by employing a nonparametric approach. Specifically, we used generalized kernel estimation to examine the rate of return to education from 1940-2005 using U.S. Census data. Our results uncovered significant heterogeneity both across and within groups. Many of the results found (on average) were in line with previous studies. For instance, we found the rate of return to education to be larger for blacks than for whites. However, we also found inconsistencies. In that same example, we found that the difference in the rate of return between blacks and whites (on average) was much larger than previously thought. Finally, we uncovered that a significant portion of the population had very low or negative rates of returns to education. We attempted to uncover which groups were more likely to possess this newfound phenomenon. At the same time we tried to find which characteristics

were most common to those with the highest and lowest rates of return.

Overall, we cannot place much emphasis on any particular point estimate from this study. The rates of return are likely impacted by omitted variables, measurement error and other econometric and economic issues. Although our results are somewhat raw, they do provide a springboard for further studies on the estimation of rates of return to education. We suggest that when estimating Mincer type regressions that authors focus on both heterogeneity across and within groups. Further, we suggest that future studies also focus on macroeconomic variables which are likely to have an affect on the perceived rate of return to education.

The Mincer regression case studied here is a specific case of the general problem of parameter heterogeneity. Whereas we believe that most economists believe that parameter heterogeneity exists, it is often ignored in economic papers. We hope that this paper highlights the problem sufficiently to encourage authors to think more about this issue and how ignoring parameter heterogeneity could be detrimental when issuing policy based on average behavior.

References

- [1] Aitchison J, Aitken CBB 1976. Multivariate Binary Discrimination by Kernel Method. *Biometrika* 63: 413-20.
- [2] Angrist J, Chen S. 2007. Long-Term Effects of Vietnam-Era Conscriptioin: Schooling, Experience and Earnings. Center for Economic Studies, U.S. Census Bureau Working Paper.
- [3] Becker GS. 1964. *Human Capital, A Theoretical and Empirical Analysis, with Special Reference to Education*. Columbia University Press: New York.
- [4] Boivin J, Giannoni MP, Mihov I. 2009. Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data. *American Economic Review* 99: 350-384.
- [5] Brand JE, Yu Xie. 2010. Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education. *American Sociological Review* 75: 273-302.
- [6] Card D. 1995. Earnings, Schooling, and Ability Revisited. *Research in Labor Economics* 14: 23-48.
- [7] Card D. 2001. Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica* 69: 1127-1160.
- [8] Card D. 2008. Returns to Schooling. *The New Palgrave Dictionary of Economics* Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan. 23 October 2010
- [9] Card D, Krueger A. 1992. School Quality and Black-White Relative Wages: A Direct Assessment. *Quarterly Journal of Economics* 107: 151-200.
- [10] Card D, Lemieux T. 2001. Can Falling Supply Explain the Rising Return to College for Younger Males? *Quarterly Journal of Economics* 116: 705-746.
- [11] Chiappori PA, Ekeland I. 2009. The Microeconomics Of Efficient Group Behavior: Identification. *Econometrica* 77: 763-799.
- [12] Chiswick B. 1978. The Effects of Americanization on the Earnings of the Foreign-born. *Journal of Political Economy* 86: 897-922.

- [13] Chiswick B, Miller P. 2004. Language Skills and Earnings among Legalized Aliens. in K. Zimmermann and A. Constant (eds.), *How Labor Migrants Fare*. Springer-Verlag: Berlin.
- [14] Christiano L, Eichenbaum M, Evans C. 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy* 113:1-45.
- [15] Clark RL, Mitchell OS. 2005. *Reinventing the Retirement Paradigm*. Oxford University Press: Oxford.
- [16] Fan J, Gijbels I. 1992. Variable Bandwidth and Local Linear Regression Smoothers. *Annals of Statistics* 20: 2008-2036.
- [17] Freeman RB. 1977. The Decline in the Economic Rewards to College Education. *Review of Economics and Statistics* 59: 18-29.
- [18] Friedberg RM 2000. You Can't Take it with You? Immigrant Assimilation and the Portability of Human Capital. *Journal of Labour Economics* 18: 221-251.
- [19] Goldin C, Katz LF. 1999. The Returns to Skill in the United States across the Twentieth Century. NBER Working Papers No. 7126.
- [20] Goldin C, Katz LF. 2008. *The Race between Education and Technology*. Belknap Press of Harvard University: Cambridge.
- [21] Harmon C, Hogan V, Walker I. 2003. Dispersion in the Economic Return to Schooling. *Labour Economics* 10: 205-214.
- [22] Heckman J, Lochner L, Todd P. 2003. Fifty years of Mincer earnings regressions. Technical Report 9732, National Bureau of Economic Research.
- [23] Heckman J, Lochner L, Todd P. 2006. Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. in E. A. Hanushek and F. Welch (eds.) *Handbook of the Economics of Education*, Volume 1. Elsevier Press: Amsterdam.
- [24] Heckman J, Lochner L, Todd P. 2008. Earnings Functions and Rates of Return. *Journal of Human Capital* 2: 1-31.
- [25] Heckman J, Polachek S. 1974. Empirical Evidence on Functional Form of the Earnings-Schooling Relationship," *Journal of the American Statistical Association* 69: 350-354.

- [26] Heckman, JJ, Schmieder DA, Urzua SS. 2009. Testing the Correlated Random Coefficient Model. NBER Working Paper Number 15463
- [27] Hildreth C, Houck J. 1968. Some Estimators for a Linear model with Random Coefficients. *Journal of the American Statistical Association* 63: 584-595.
- [28] Hsiao C, Li Q, Racine J. 2007. A Consistent Model Specification Test with Mixed Discrete and Continuous Data. *Journal of Econometrics* 140: 802-826.
- [29] Hurvich CM, Simonoff JS, Tsai CL. 1998. Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society, Series B* 60: 271-293.
- [30] Imbens G, Angrist J. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica* 62: 467-476.
- [31] Jaimovich N, Rebelo S. 2009 Can News about the Future Drive the Business Cycle? *American Economic Review* 99:1097-1118.
- [32] Kniesner TJ, Padilla AH, Polachek SW. 1978. The Rate of Return to Schooling and the Business Cycle. *Journal of Human Resources* 13: 265-277.
- [33] Koop G, Tobias J. 2004. Learning About Heterogeneity in Returns to Schooling. *Journal of Applied Econometrics* 19: 827-849.
- [34] Lazear E. 1977. Schooling as a Wage Depressant. *Journal of Human Resources* 12: 164-76.
- [35] Lemieux T. and D. Card. 2001. Education, Earnings, and the ‘Canadian G.I. Bill’. *Canadian Journal of Economics* 34(2): 313-344.
- [36] Li Q, Racine J. 2004. Cross-Validated Local Linear Nonparametric Regression. *Statistica Sinica* 14: 485-512.
- [37] Li Q, Racine J. 2007. *Nonparametric Econometrics: Theory and Practice*. Princeton University Press: Princeton.
- [38] Martins, P.S. and P.T. Pereira. 2004. Does Education Reduce Wage Inequality? Quantile Regression Evidence From 16 Countries. *Labour Economics* 11:355–371.

- [39] Mincer J. 1974. *Schooling, Experience, and Earnings*. Columbia University Press: New York.
- [40] Pagan A, Ullah A. 1999. *Nonparametric Econometrics*. Cambridge University Press: Cambridge.
- [41] Park JH 1994. Estimation of Sheepskin Effects and Returns to Schooling using the Old and the New CPS Measures of Educational Attainment. Working Paper #338, Industrial Relations Section, Princeton University.
- [42] Parry IWH, Small KA. 2009. Should Urban Transit Subsidies Be Reduced? *American Economic Review* 99: 700-724.
- [43] Pesaran H. 2006. Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica* 74: 967-1012
- [44] Pischke J, von Wachter T. 2008. Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation. *Review of Economics and Statistics* 90: 592-598
- [45] Polachek SW. 2008. Earnings Over the Lifecycle: The Mincer Earnings Function and Its Applications. *Foundations and Trends in Microeconomics* 4: 165-272.
- [46] Polachek SW, Kim M-K. 1994. Panel Estimates of the Gender Wage Gap: Individual-Specific Intercept and Individual Specific Slope Models. *Journal of Econometrics* 61: 23-42.
- [47] Racine J, Li Q. 2004. Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data. *Journal of Econometrics* 119: 99-130.
- [48] Rau-Binder, Tomas. 2006. Semiparametric Estimation of Microeconomic Models with Engogenous Regressors and Sorting. *University of California Berkeley Department of Economics Unpublished Manuscript*.
- [49] Rivera-Batiz F. 2007. How Do Migrants from Latin America and the Caribbean Fare in the US Labour Market? *World Economy* 30: 1399-1429.
- [50] Rubin H. 1950. Note on Random Coefficients. in T Koopmans *Statistical Inference in Dynamic Economic Models* (NY and Londaon: John Wiley & Sons, Inc.) 419-421.
- [51] Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall: London.

- [52] Smets F, Wouters R. 2003. An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. *Journal of the European Economic Association* 1:1123-75.
- [53] Soldo BJ, Mitchell OS, Tfray R, McCabe JF. 2006. Cross-Cohort Differences in Health on the Verge of Retirement. NBER Working Paper No. 12762.
- [54] Welch F. 1973. Black and White Differences in Returns to Schooling. *American Economic Review* 63: 893-907.

Table 1: Baseline Results for Schooling Rate of Return

Panel A: Parametric Results									
Year	All	20-29	30-39	40-49	50-59	White	Black	Immigrant	Native
1940	0.129	0.188	0.143	0.108	0.081	0.124	0.103	0.061	0.149
1950	0.084	0.106	0.091	0.079	0.061	0.071	0.085	0.049	0.087
1960	0.091	0.131	0.095	0.084	0.076	0.081	0.093	0.062	0.094
1970	0.083	0.125	0.092	0.085	0.074	0.077	0.090	0.071	0.085
1980	0.080	0.098	0.093	0.078	0.057	0.074	0.101	0.074	0.076
1990	0.107	0.166	0.112	0.102	0.064	0.104	0.132	0.088	0.111
2000	0.108	0.158	0.119	0.090	0.068	0.105	0.135	0.092	0.110
2005	0.123	0.191	0.133	0.117	0.082	0.118	0.159	0.092	0.130

Panel B.1: Nonparametric Results (Median Values)									
Year	All	20-29	30-39	40-49	50-59	White	Black	Immigrant	Native
1940	0.151	0.201	0.144	0.118	0.107	0.148	0.110	0.078	0.176
1950	0.082	0.118	0.092	0.090	0.070	0.078	0.097	0.042	0.091
1960	0.093	0.144	0.089	0.085	0.080	0.091	0.093	0.058	0.097
1970	0.094	0.108	0.112	0.084	0.073	0.088	0.098	0.083	0.084
1980	0.096	0.086	0.110	0.084	0.060	0.090	0.142	0.082	0.098
1990	0.133	0.149	0.141	0.136	0.097	0.123	0.205	0.110	0.127
2000	0.135	0.137	0.166	0.113	0.070	0.124	0.190	0.114	0.144
2005	0.143	0.163	0.162	0.152	0.090	0.139	0.239	0.124	0.142

Panel B.2: Nonparametric Results (Coefficients of Variation)									
1940	0.281	0.316	0.158	0.299	0.582	0.411	0.486	0.631	0.312
1950	0.872	1.141	0.665	0.323	0.240	0.628	0.802	4.228	0.387
1960	0.787	0.558	0.365	0.419	0.393	1.010	0.305	1.209	0.597
1970	0.648	0.882	0.320	0.123	0.073	0.833	0.463	1.152	0.653
1980	0.629	1.131	0.288	0.254	0.234	1.111	0.540	0.191	0.916
1990	0.316	1.033	0.275	0.421	0.326	0.698	0.395	0.369	0.242
2000	0.413	0.744	0.305	0.246	0.184	0.294	0.480	0.414	0.536
2005	0.373	0.899	0.191	0.314	0.242	0.375	0.446	0.456	0.376

¹ Notes: The results in Panel A are computed from IPUMS data using Mincer (1974) specification. See text for details.

² The results in Panels B.1 and B.2 are computed from IPUMS data using a nonparametric specification. Median rate of return values reported for each indicated group. See text for details.

Table 2: Schooling Rate of Return Parametric Estimates by Race

Panel A: Parametric Results								
	Black							
Age Group	1940	1950	1960	1970	1980	1990	2000	2005
20-29	0.134	0.131	0.161	0.157	0.160	0.228	0.210	0.266
30-39	0.108	0.095	0.097	0.092	0.117	0.139	0.149	0.175
40-49	0.089	0.071	0.074	0.084	0.086	0.118	0.112	0.138
50-59	0.082	0.063	0.072	0.077	0.049	0.073	0.095	0.099
	White							
	1940	1950	1960	1970	1980	1990	2000	2005
20-29	0.191	0.091	0.124	0.119	0.089	0.163	0.162	0.201
30-39	0.136	0.079	0.083	0.091	0.083	0.102	0.123	0.140
40-49	0.095	0.066	0.078	0.081	0.073	0.103	0.093	0.115
50-59	0.071	0.050	0.064	0.063	0.054	0.058	0.070	0.065
Panel B: Nonparametric Results								
	Black							
Age Group	1940	1950	1960	1970	1980	1990	2000	2005
20-29	0.130	0.122	0.146	0.174	0.194	0.267	0.184	0.265
30-39	0.111	0.097	0.098	0.098	0.164	0.227	0.242	0.299
40-49	0.100	0.078	0.078	0.072	0.107	0.163	0.199	0.256
50-59	0.077	0.069	0.072	0.069	0.056	0.112	0.144	0.136
	White							
	1940	1950	1960	1970	1980	1990	2000	2005
20-29	0.190	0.092	0.128	0.087	0.066	0.117	0.150	0.209
30-39	0.139	0.083	0.084	0.100	0.128	0.130	0.140	0.150
40-49	0.129	0.070	0.085	0.091	0.096	0.111	0.109	0.127
50-59	0.084	0.055	0.083	0.063	0.064	0.160	0.081	0.084

¹ Notes: The results in Panel A are computed from IPUMS data using Mincer (1974) specification. See text for details.

² The results in Panel B are computed from IPUMS data using a nonparametric specification. Median rate of return values reported for each indicated group. See text for details.

Table 3: Schooling Rate of Return Parametric Estimates by Immigrant Status

Panel A: Parametric Results								
	Immigrant							
	1940	1950	1960	1970	1980	1990	2000	2005
20-29	0.141	0.087	0.082	0.117	0.089	0.117	0.142	0.139
30-39	0.086	0.078	0.078	0.077	0.082	0.103	0.108	0.111
40-49	0.052	0.057	0.066	0.067	0.064	0.082	0.082	0.090
50-59	0.039	0.021	0.052	0.054	0.057	0.075	0.066	0.074
	Native							
	1940	1950	1960	1970	1980	1990	2000	2005
20-29	0.195	0.103	0.132	0.131	0.105	0.166	0.173	0.211
30-39	0.149	0.092	0.097	0.092	0.087	0.123	0.131	0.142
40-49	0.140	0.083	0.088	0.091	0.075	0.107	0.101	0.126
50-59	0.107	0.075	0.078	0.078	0.061	0.070	0.058	0.082
Panel B: Nonparametric Results								
	Immigrant							
	1940	1950	1960	1970	1980	1990	2000	2005
20-29	0.129	0.036	0.075	0.089	0.087	0.107	0.146	0.149
30-39	0.099	0.035	0.072	0.080	0.090	0.132	0.137	0.145
40-49	0.066	0.087	0.055	0.086	0.072	0.103	0.112	0.116
50-59	0.057	0.017	0.050	0.077	0.056	0.091	0.091	0.103
	Native							
	1940	1950	1960	1970	1980	1990	2000	2005
20-29	0.200	0.082	0.130	0.089	0.100	0.163	0.151	0.219
30-39	0.166	0.086	0.089	0.082	0.100	0.132	0.167	0.159
40-49	0.157	0.095	0.079	0.092	0.106	0.112	0.141	0.130
50-59	0.149	0.092	0.108	0.069	0.075	0.085	0.110	0.089

¹ Notes: The results in Panel A are computed from IPUMS data using Mincer (1974) specification. See text for details.

² The results in Panel B are computed from IPUMS data using a non-parametric specification. Median rate of return values reported for each indicated group. See text for details.

Table 4: Schooling Rate of Return Estimates by Education Level

Panel A: Parametric Results				
Year	HS & Below	High School	Some College	College & Above
1940	0.131	0.132	0.127	0.126
1950	0.091	0.094	0.078	0.080
1960	0.100	0.100	0.091	0.089
1970	0.089	0.090	0.076	0.087
1980	0.089	0.093	0.084	0.084
1990	0.110	0.104	0.105	0.103
2000	0.108	0.100	0.103	0.103
2005	0.122	0.113	0.116	0.117

Panel B.1: Nonparametric Results (Median Values)

Year	HS & Below	High School	Some College	College & Above
1940	0.153	0.147	0.128	0.125
1950	0.128	0.059	0.043	-0.007
1960	0.101	0.091	0.082	0.011
1970	0.086	0.098	0.087	0.048
1980	0.110	0.116	0.071	0.042
1990	0.125	0.150	0.135	0.092
2000	0.129	0.172	0.137	0.056
2005	0.109	0.163	0.167	0.099

Panel B.2: Nonparametric Results (Coefficients of Variation)

Year	HS & Below	High School	Some College	College & Above
1940	0.2884	0.2553	0.2677	0.2538
1950	0.5114	0.4853	8.3435	41.8641
1960	0.7715	0.3764	0.8224	7.6684
1970	0.7071	0.4337	0.5383	1.1526
1980	0.7476	0.2924	0.599	1.4072
1990	0.4319	0.241	0.2595	0.3478
2000	0.5749	0.2918	0.2504	0.7711
2005	0.497	0.267	0.3011	0.5204

¹ Notes: The results in Panel A are computed from IPUMS data using Mincer (1974) specification. See text for details.

² The results in Panel B are computed from IPUMS data using a nonparametric specification. Median rate of return values reported for each indicated group. See text for details.

Table 5. Percentage Difference in Attributes for Top 5% from the Remainder of the Sample

Year	Variable	Age 20-29	Age 30-39	Age 40-49	Age 50-59	White	Black	Immigrant	Native
1940	Married	0.158	-0.025	-0.002	-0.007	-0.500	-0.341	-0.331	-0.390
	Child < 5	0.132	-0.019	0.116	0.014	-0.094	-0.008	0.107	-0.050
	Black	0.022	0.095	0.161	-0.018			0.009	0.016
	Immigrant	0.011	0.038	0.048	-0.033	-0.109	-0.009		
	Age 20-29					0.697	0.623	0.789	0.667
	Age 30-39					-0.296	-0.284	-0.094	-0.315
	Age 40-49					-0.240	-0.220	-0.366	-0.221
	Age 50-59					-0.161	-0.119	-0.330	-0.131
	America							0.167	
	Europe & Oceania							-0.189	
	Asia							0.011	
1950	Married	0.049	0.014	-0.013	-0.009	-0.243	-0.236	-0.069	-0.222
	Child < 5	-0.025	-0.082	0.010	0.041	0.025	-0.047	0.101	0.079
	Black	-0.052	0.092	0.082	0.081			0.000	0.131
	Immigrant	0.006	0.036	0.041	0.027	-0.027	-0.007		
	Age 20-29					0.627	0.411	0.212	0.710
	Age 30-39					-0.249	-0.202	0.020	-0.311
	Age 40-49					-0.224	-0.247	-0.037	-0.242
	Age 50-59					-0.153	0.038	-0.195	-0.157
	America							0.113	
	Europe & Oceania							-0.176	
	Asia							0.041	
1960	Married	-0.183	-0.031	-0.031	0.001	-0.171	-0.491	-0.205	-0.209
	Child < 5	-0.173	0.048	-0.012	0.038	0.085	-0.084	0.024	0.077
	Black	0.050	0.145	0.164	0.188			0.002	0.062
	Immigrant	0.013	0.027	0.020	0.031	0.002	-0.009		
	Age 20-29					0.334	0.765	0.430	0.588
	Age 30-39					-0.051	-0.317	-0.108	-0.242
	Age 40-49					-0.178	-0.263	-0.116	-0.197
	Age 50-59					-0.105	-0.184	-0.206	-0.149
	America							0.037	
	Europe & Oceania							-0.170	
	Asia							0.037	
1970	Married	-0.188	-0.010	-0.004	-0.024	-0.305	-0.357	-0.070	-0.269
	Child < 5	-0.235	0.002	-0.009	0.013	-0.019	0.057	0.165	0.044
	Black	-0.026	0.025	0.048	0.053			-0.015	-0.012
	Immigrant	0.015	0.046	0.041	0.068	-0.009	-0.014		

Table 5 – (cont.) Percentage Difference in Attributes for Top 5% from the Remainder of the Sample

Year	Variable	Age 20-29	Age 30-39	Age 40-49	Age 50-59	White	Black	Immigrant	Native
	Age 20-29					0.654	0.725	0.078	0.706
	Age 30-39					-0.220	-0.281	0.365	-0.237
	Age 40-49					-0.254	-0.254	-0.212	-0.263
	Age 50-59					-0.180	-0.191	-0.231	-0.205
	America							-0.131	
	Europe & Oceania							-0.100	
	Asia							0.292	
1980	Married	-0.210	0.042	-0.001	0.004	-0.321	-0.381	-0.178	-0.283
	Child < 5	-0.184	-0.085	0.042	0.004	-0.012	-0.016	0.049	-0.020
	Black	-0.021	0.054	0.066	-0.011			-0.012	0.008
	Immigrant	0.016	0.006	0.146	-0.002	0.000	0.001		
	Age 20-29					0.610	0.586	0.181	0.402
	Age 30-39					-0.251	-0.259	0.218	-0.110
	Age 40-49					-0.185	-0.181	-0.230	-0.172
	Age 50-59					-0.174	-0.146	-0.169	-0.120
	America							-0.199	
	Europe & Oceania							-0.062	
	Asia							0.281	
1990	Married	-0.127	0.055	0.023	0.003	-0.123	-0.320	-0.122	-0.595
	Child < 5	-0.052	-0.110	-0.064	0.002	-0.012	0.040	-0.071	-0.161
	Black	0.017	0.040	0.008	0.008			0.016	-0.014
	Immigrant	-0.011	-0.029	-0.045	-0.034	0.002	-0.024		
	Age 20-29					0.312	0.677	0.090	0.735
	Age 30-39					-0.238	-0.317	-0.206	-0.335
	Age 40-49					0.060	-0.226	0.262	-0.250
	Age 50-59					-0.134	-0.134	-0.146	-0.150
	America							-0.069	
	Europe & Oceania							0.035	
	Asia							0.019	
2000	Married	-0.128	0.018	0.006	-0.001	-0.493	-0.162	-0.549	-0.062
	Child < 5	-0.025	-0.179	-0.046	0.011	-0.098	0.035	-0.183	0.070
	Black	0.019	0.035	0.037	0.043	0.000	0.000	0.012	0.043
	Immigrant	-0.027	-0.032	-0.002	0.127	-0.012	-0.036		
	Age 20-29					0.759	0.036	0.782	-0.061
	Age 30-39					-0.260	0.288	-0.355	0.483
	Age 40-49					-0.299	-0.174	-0.267	-0.219
	Age 50-59					-0.201	-0.150	-0.161	-0.202
	America							-0.186	

Table 5 – (cont.) Percentage Difference in Attributes for Top 5% from the Remainder of the Sample

Year	Variable	Age 20-29	Age 30-39	Age 40-49	Age 50-59	White	Black	Immigrant	Native
	Europe & Oceania							0.018	
	Asia							0.152	
2005	Married	-0.137	-0.031	-0.003	-0.007	-0.660	-0.281	-0.402	-0.689
	Child < 5	-0.072	-0.021	-0.087	-0.003	-0.143	0.040	-0.063	-0.134
	Black	0.010	0.017	0.025	0.042	0.000	0.000	0.013	0.008
	Immigrant	-0.037	-0.020	-0.052	-0.013	-0.040	-0.049		
	Age 20-29					0.828	0.363	0.726	0.827
	Age 30-39					-0.264	0.116	-0.230	-0.262
	Age 40-49					-0.306	-0.270	-0.300	-0.303
	Age 50-59					-0.259	-0.209	-0.196	-0.262
	America							-0.184	
	Europe & Oceania							0.029	
	Asia							0.144	

Notes: Each element represents the percentage difference for the bottom 5% of people in the column group who have the row attribute from the top 95% of people in the column group who have the row attribute. Specifically, we take the average of the characteristic (the percentage of people in that group who have that characteristic) for the people in the top 5% and subtract the average of the characteristic for the people in the bottom 95% and test the significance (t-test) of that difference.

Table 6. Percentage Difference in Attributes for Bottom 5% from the Remainder of the Sample

Year	Variable	Age 20-29	Age 30-39	Age 40-49	Age 50-59	White	Black	Immigrant	Native
1940	Married	0.139	0.003	-0.013	-0.007	0.161	0.059	0.062	0.192
	Child < 5	0.063	0.041	0.053	0.027	-0.087	-0.043	-0.107	-0.107
	Black	0.226	0.086	0.106	0.078			-0.006	0.017
	Immigrant	0.012	0.038	0.271	0.366	0.306	0.006		
	Age 20-29					-0.330	-0.144	-0.101	-0.353
	Age 30-39					-0.247	-0.206	-0.189	-0.198
	Age 40-49					0.237	0.328	-0.281	0.297
	Age 50-59					0.341	0.022	0.571	0.254
	America							-0.052	
	Europe & Oceania							0.061	
	Asia							-0.005	
1950	Married	-0.211	-0.006	0.005	0.021	-0.372	-0.045	-0.052	-0.391
	Child < 5	-0.203	0.059	0.008	-0.009	-0.195	-0.103	-0.025	-0.167
	Black	-0.038	0.034	0.080	-0.051			-0.004	-0.058
	Immigrant	-0.003	0.000	0.048	-0.033	0.021	0.007		
	Age 20-29					0.441	-0.071	0.109	0.568
	Age 30-39					-0.305	-0.156	-0.059	-0.319
	Age 40-49					-0.177	0.029	-0.089	-0.165
	Age 50-59					0.041	0.198	0.039	-0.083
	America							0.003	
	Europe & Oceania							0.001	
	Asia							-0.010	
1960	Married	-0.195	-0.010	0.002	-0.013	-0.100	0.103	-0.063	-0.071
	Child < 5	-0.224	0.090	0.042	0.001	-0.080	-0.270	-0.020	-0.082
	Black	-0.051	-0.049	-0.018	0.005			-0.007	0.007
	Immigrant	-0.004	0.030	0.022	0.034	0.035	0.017		
	Age 20-29					0.069	-0.289	0.074	0.045
	Age 30-39					-0.097	-0.317	-0.100	-0.107
	Age 40-49					-0.098	-0.226	-0.046	-0.017
	Age 50-59					0.126	0.832	0.073	0.079
	America							0.042	
	Europe & Oceania							-0.052	
	Asia							0.031	
1970	Married	-0.404	-0.030	-0.009	-0.002	-0.344	0.094	-0.161	-0.291
	Child < 5	-0.298	0.048	0.101	0.001	-0.141	-0.202	-0.076	-0.095
	Black	-0.008	0.003	0.039	0.009			-0.019	0.042
	Immigrant	0.005	0.063	0.043	-0.013	0.018	0.006		

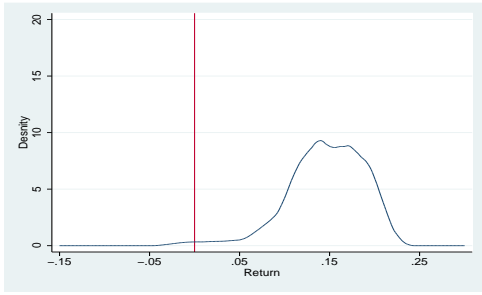
Table 6 – (cont.) Percentage Difference in Attributes for Top 5% from the Remainder of the Sample

Year	Variable	Age 20-29	Age 30-39	Age 40-49	Age 50-59	White	Black	Immigrant	Native
	Age 20-29					0.416	-0.296	0.212	0.404
	Age 30-39					-0.250	-0.232	-0.150	-0.225
	Age 40-49					-0.206	-0.033	-0.111	-0.071
	Age 50-59					0.040	0.561	0.049	-0.107
	America							0.034	
	Europe & Oceania							-0.056	
	Asia							0.011	
1980	Married	-0.353	-0.062	-0.007	-0.011	-0.444	0.199	0.186	-0.260
	Child < 5	-0.210	0.080	0.037	0.004	-0.123	-0.168	-0.158	-0.157
	Black	-0.008	-0.029	-0.039	-0.035			-0.018	0.003
	Immigrant	0.005	0.070	0.078	0.038	0.006	0.016		
	Age 20-29					0.484	-0.382	-0.318	0.159
	Age 30-39					-0.258	-0.168	-0.285	-0.254
	Age 40-49					-0.179	-0.036	0.242	0.050
	Age 50-59					-0.048	0.585	0.361	0.045
	America							0.053	
	Europe & Oceania							0.040	
	Asia							-0.091	
1990	Married	-0.299	-0.034	0.018	0.003	-0.316	0.207	0.107	0.202
	Child < 5	-0.159	0.097	0.023	0.002	-0.115	-0.117	-0.031	-0.168
	Black	0.033	-0.034	-0.023	-0.020			-0.033	0.005
	Immigrant	-0.030	0.176	0.115	0.069	0.021	0.054		
	Age 20-29					0.440	-0.327	-0.155	-0.305
	Age 30-39					-0.331	-0.182	-0.098	-0.293
	Age 40-49					-0.213	-0.055	0.058	-0.088
	Age 50-59					0.104	0.564	0.194	0.686
	America							0.195	
	Europe & Oceania							-0.054	
	Asia							-0.106	
2000	Married	-0.262	0.056	0.032	0.025	0.179	0.072	0.128	-0.090
	Child < 5	-0.120	0.107	0.101	0.003	-0.061	-0.054	-0.004	-0.114
	Black	0.008	-0.049	-0.043	-0.034	0.000	0.000	-0.042	-0.016
	Immigrant	-0.010	0.387	0.155	0.037	0.055	0.110		
	Age 20-29					-0.231	-0.179	-0.215	0.110
	Age 30-39					-0.177	-0.074	0.002	-0.287
	Age 40-49					0.082	0.126	0.126	-0.152
	Age 50-59					0.327	0.127	0.087	0.329
	America							0.185	

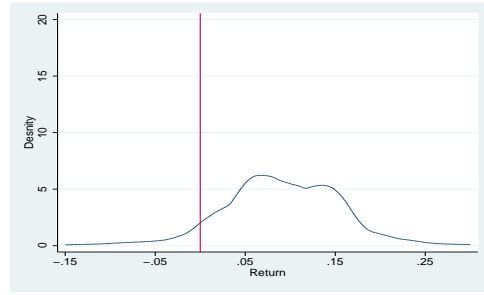
Table 6 – (cont.) Percentage Difference in Attributes for Top 5% from the Remainder of the Sample

Year	Variable	Age 20-29	Age 30-39	Age 40-49	Age 50-59	White	Black	Immigrant	Native
	Europe & Oceania							-0.053	
	Asia							-0.093	
2005	Married	-0.176	0.047	0.076	0.055	0.191	0.189	0.168	0.218
	Child < 5	-0.073	0.104	0.071	0.006	0.007	-0.005	-0.035	-0.034
	Black	0.017	-0.047	-0.043	-0.034			-0.037	-0.043
	Immigrant	0.054	0.281	0.041	0.009	0.078	0.168		
	Age 20-29					-0.208	-0.221	-0.214	-0.230
	Age 30-39					-0.086	-0.127	-0.084	-0.135
	Age 40-49					0.032	0.078	0.032	-0.030
	Age 50-59					0.262	0.269	0.266	0.394
	America							-0.012	
	Europe & Oceania							0.013	
	Asia							0.016	

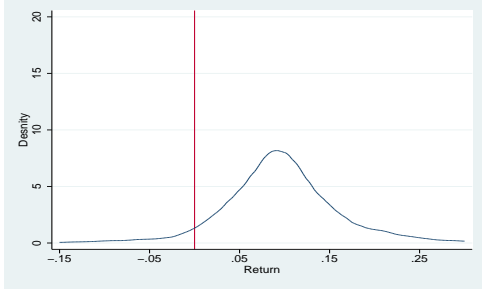
Notes: Each element represents the percentage difference for the bottom 5% of people in the column group who have the row attribute from the top 95% of people in the column group who have the row attribute. Specifically, we take the average of the characteristic (the percentage of people in that group who have that characteristic) for the people in the bottom 5% and subtract the average of the characteristic for the people in the top 95% and test the significance (t-test) of that difference.



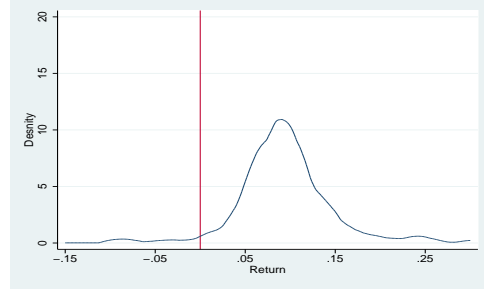
(a) 1940



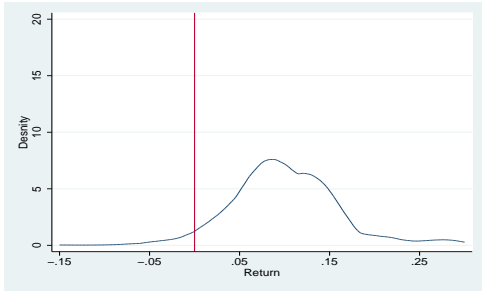
(b) 1950



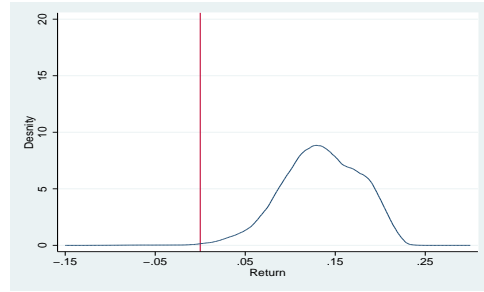
(c) 1960



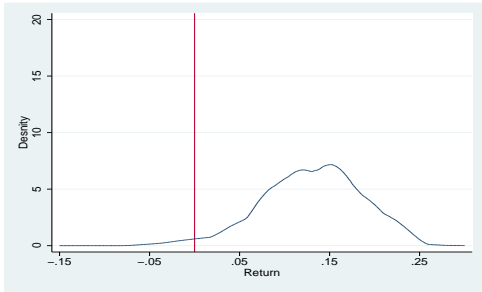
(d) 1970



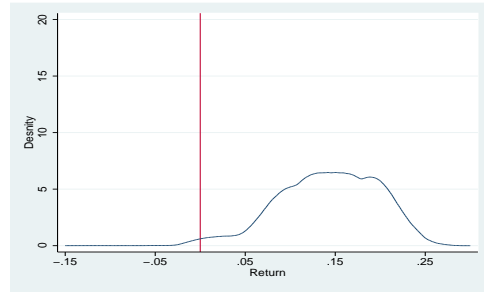
(e) 1980



(f) 1990

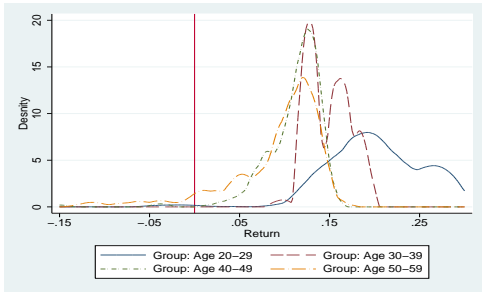


(g) 2000

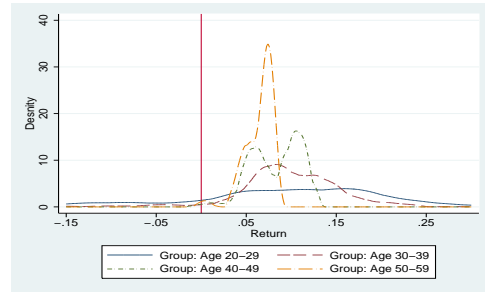


(h) 2005

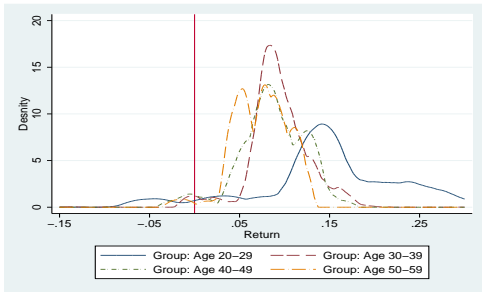
Figure 1: Nonparametric Results by Year



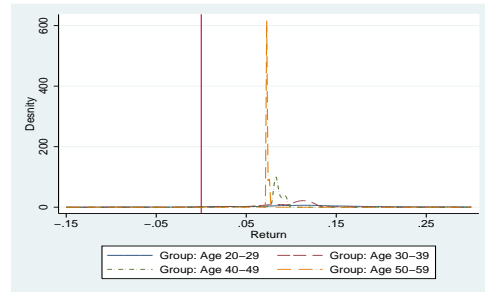
(a) 1940



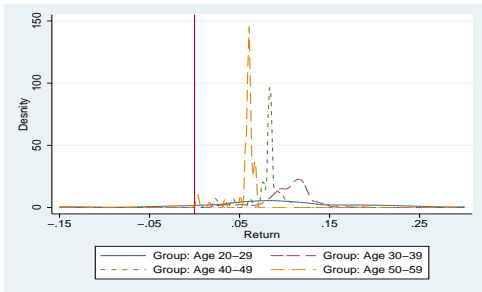
(b) 1950



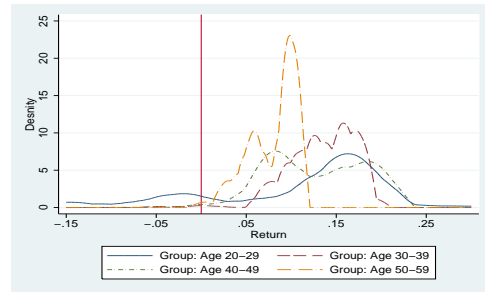
(c) 1960



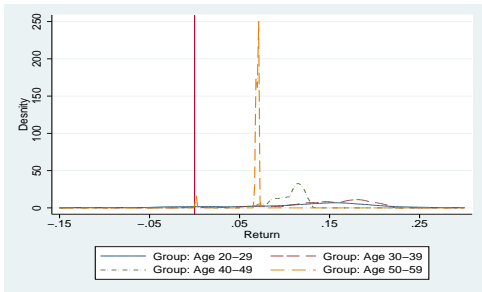
(d) 1970



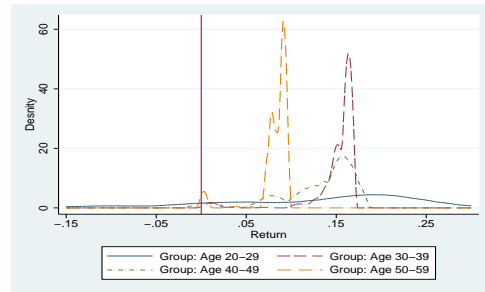
(e) 1980



(f) 1990

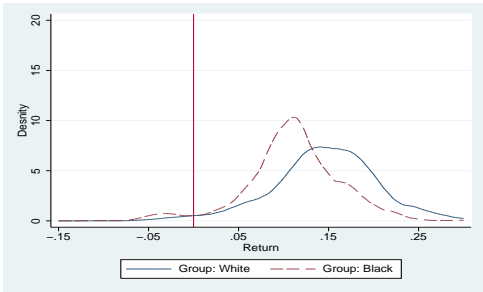


(g) 2000

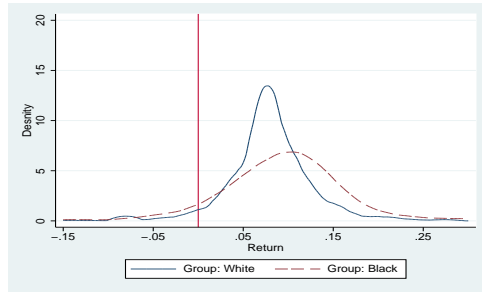


(h) 2005

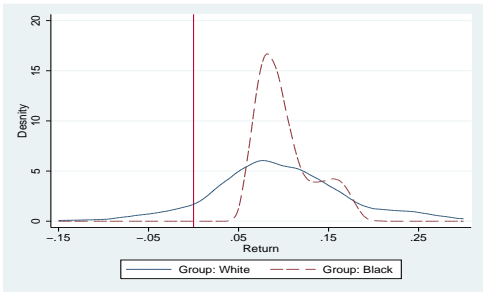
Figure 2: Nonparametric Results by Year and Age Groups



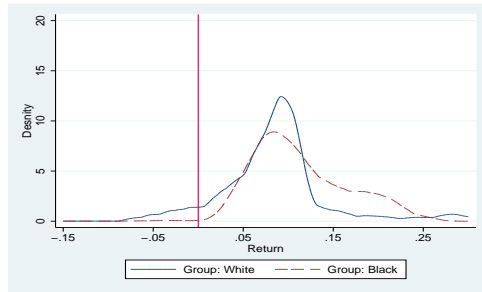
(a) 1940



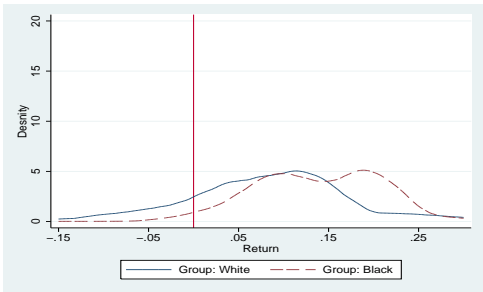
(b) 1950



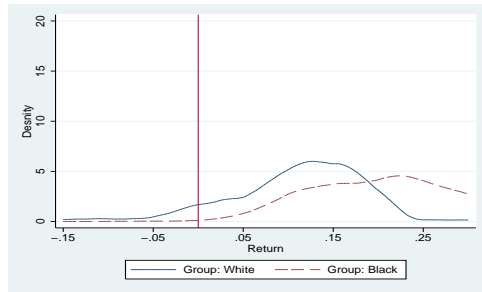
(c) 1960



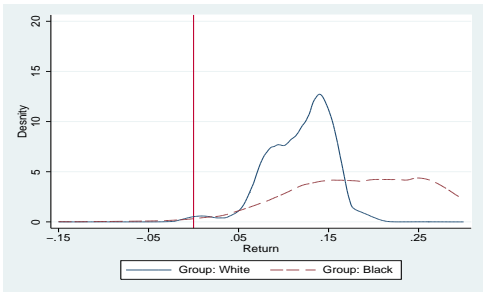
(d) 1970



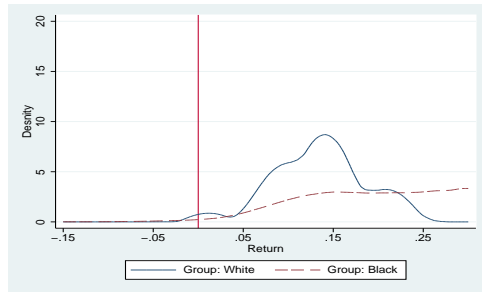
(e) 1980



(f) 1990

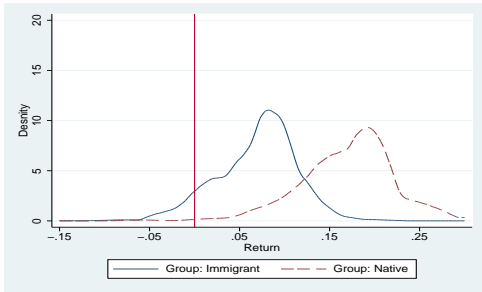


(g) 2000

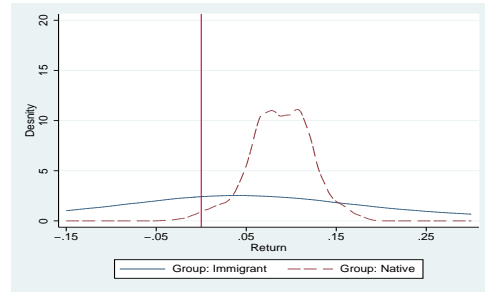


(h) 2005

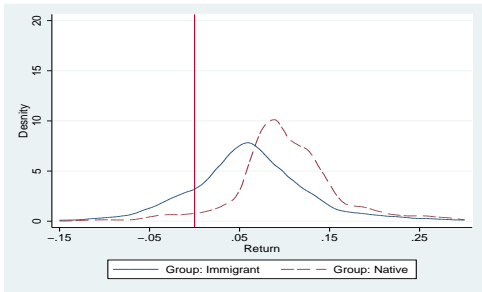
Figure 3: Nonparametric Results by Year and Race



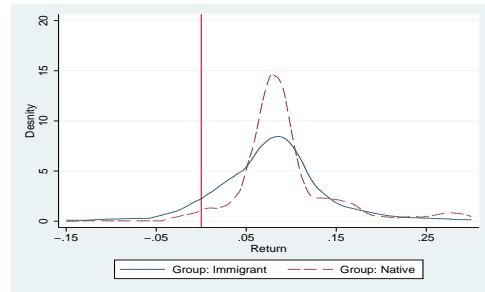
(a) 1940



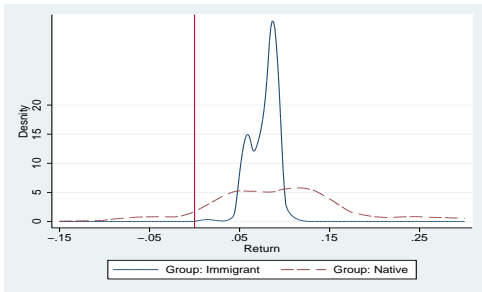
(b) 1950



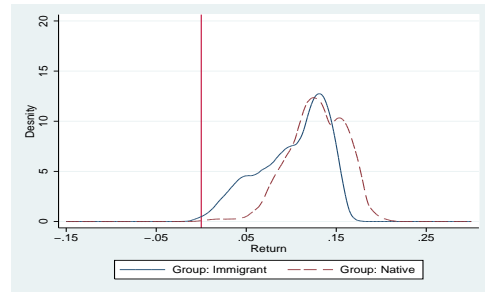
(c) 1960



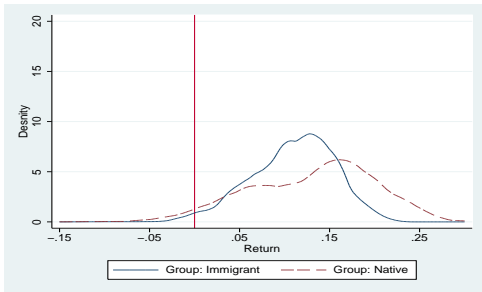
(d) 1970



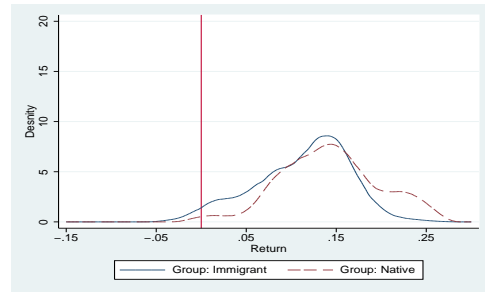
(e) 1980



(f) 1990



(g) 2000



(h) 2005

Figure 4: Nonparametric Results by Year and Immigration Status