NBER WORKING PAPER SERIES

THE SCHOOLING QUALITY-EARNINGS
RELATIONSHIP: USING ECONOMIC
THEORY TO INTERPRET FUNCTIONAL
FORMS CONSISTENT WITH THE EVIDENCE

James Heckman
Anne Layne-Farrar
Petra Todd

THE SCHOOLING QUALITY-EARNINGS
RELATIONSHIP: USING ECONOMIC
THEORY TO INTERPRET FUNCTIONAL
FORMS CONSISTENT WITH THE EVIDENCE

ABSTRACT

This paper investigates the economic and empirical foundations of the evidence relating earnings to schooling quality. We replicate the Card-Krueger model for Census years 1970, 1980 and 1990 and find that it consistently produces a strong relationship between schooling quality and the rate of return to schooling. We test key identifying assumptions used by Card and Krueger and others. Several assumptions are rejected. When they are relaxed, the evidence for a strong effect of schooling quality on earning is greatly weakened. A crucial identifying assumption is the absence of selective migration on the basis of earnings. Nonparametric tests strongly reject this hypothesis. The conventional assumption of linearity of the earnings-schooling relationship widely used in the literature is also rejected. The only surviving evidence of any schooling quality effect is in the return to college education. We also test and reject conventional efficiency unit models of the pricing of labor services. The empirically concordant model of earnings is a model of heterogeneous human capital in which regional shocks affect the prices of less-skilled labor.

James Heckman
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER

Anne Layne-Farrar
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637

Petra Todd
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637

Introduction

This paper investigates the economic and empirical foundations of the evidence relating earnings to schooling quality. Numerous papers have been written on the topic, but most have adopted one of two empirical specifications. In one model, schooling quality raises earnings by the same percentage independently of the level of schooling attained (see, e.g., Johnson and Stafford, 1973). In the second model, schooling quality raises earnings by increasing the coefficient on schooling in a log earnings-schooling relationship (Behrman and Birdsall (1983)). These specifications are largely justified on intuitive grounds. No explicit human capital accumulation processes are discussed and the forces of demand and supply are usually kept well-hidden in the background.

We examine the economic justification for commonly-used specifications of the earnings - quality relationship and present tests to discriminate among competing models. We interpret the earnings-schooling-quality relationship as a pricing equation for a market with heterogenous skills. Different economic models of heterogeneous skills produce different restrictions on earnings equations. We test these restrictions, reject a variety of conventional models and produce a new model of the quality-earnings relationship with a diminished role for quality.

We address the following six questions concerning the choice of an appropriate functional form for modeling the earnings-quality relationship.

Q1. What guidance does economic theory provide on how to introduce schooling quality into earnings equations?

Q2. Are the data consistent with the basic premise that, regardless of the model adopted, higher schooling quality as measured by indices such as pupil-teacher ratios, term length and teacher salary, increases earnings?

Q3. Under what assumptions about migration, the operation of labor markets, the pricing of skills and the choices of individuals are quality effects identified under different estimation strategies?

Q4. What are the parameters of interest in evaluating the impact of schooling quality?

Q5. How sensitive are estimates of the impact of schooling quality on earnings to alternative specifications of the functional form of the earnings equation? How sensitive are they to alternative sets of the regressors introduced to capture market forces? Which functional forms are most consistent with the data?

Q6. How might a model of the endogenous determination of schooling quality explain the empirical regularity of larger estimated effects of quality on earnings when more aggregated measures of quality are introduced into earnings equation?

This paper offers the following answers to these questions.

A1. Most previous empirical studies of the earnings-quality relationship estimate a functional form that is a variant of the standard Mincer (1974) log earnings model, introducing quality in one of two ways suggested by Behrman and Birdsall (1983). A drawback of the Mincer model in this context is that the mechanism determining the return to education is not clearly specified. The model is not interpretable as a pricing equation emerging from a general equilibrium model of factor services in which payments to factors are determined by the demand and supply for heterogeneous human capital. We present some interpretable models of schooling quality and earnings determination that explicitly incorporate demand and supply factors into the estimation of pricing equations. We find that widely-used efficiency units models of earnings do not rationalize the Census data used in our study. However, a model with heterogeneous human capital in which high skilled labor trades in a national market but low skilled labor is affected by local labor market shocks is consistent with the evidence. We introduce schooling quality into the earnings equation through a human capital transformation function.

A2. Better education improves earnings. However, the evidence is much less clear that the empirical proxies for quality actually measure the value of an individual's education. A central premise in most strategies designed to estimate the effect of measured quality on earnings is the assumption that quality is an immutable characteristic yielding the same increase in the return to education in all labor markets. This premise can be tested nonparametrically using the following insight. If individuals are educated in their state of birth, then for any education level in any given labor market, ranks of schooling

2

quality by states of birth should be correlated with ranks of wages or rates of return classified by state of birth. Such a correlation is not found for any of the quality variables widely used in the literature, raising doubts about their explanatory power. A weaker form of this test looks at whether state of birth wage rankings are stable across different labor markets. There is some support for stability of these rankings, but this evidence is consistent with the operation of any common environmental variable like family background - not just school quality - and therefore is not decisive on the question of whether empirical measures of schooling quality raise earnings.

A3. A key assumption in many of the studies using aggregated schooling quality data to explain individual earnings is that migration is not selective on the unobservables of the destination wage equation. Our empirical evidence indicates that this assumption is violated and that estimation strategies need to account for nonrandom migration. When this is done, estimated quality effects often weaken or become uninterpretable. Our evidence is consistent with different effects of quality on the return to schooling in different labor markets. This causes people to sort into different regions to realize favorable economic opportunities. More generally, we find systematic patterns of comparative advantage in the sorting of persons into regions.

A4. Several different parameters of interest have been defined in the literature. By focusing on certain parameters to the exclusion of others, most studies give an incomplete picture of the effect of quality on earnings. For example, papers by Card and Krueger (1992a,b) and Johnson and Stafford (1973) answer different questions. Card and Krueger focus on the impact of quality on the return to schooling (the slope of the earnings function) while Johnson and Stafford estimate the effect of schooling quality on intercepts in earnings equations. We examine both of these effects as well as the role of regional aggregate levels of schooling quality in determining returns to education.

A5. Estimated impacts of quality on earnings are fragile with respect to the specification of the earnings function. Models linear in schooling, commonly used in the literature on schooling quality, are rejected by the data. Nonlinear models reveal that the only impact of quality on earnings comes through the marginal increment to earnings arising from college attendance. Alternative specifications of regional aggregate demand

3

and supply equations greatly affect the estimated impact of aggregate schooling quality on regional returns. Introducing empirically concordant interactions between regions of birth and regions of residence to account for selective migration weakens the estimated within-region impact of schooling quality on the return to price of schooling. So does the introduction of an empirical model that allows the effect of increases in quality on returns to education to differ across regions. Empirical estimates of the impact of educational quality on returns to education earnings are not robust to alternative specifications of functional forms nor to the inclusion of additional regressors, although the additional regressors and alternative functional forms are consistent with the data.

A6. A model of collective decision-making about school funding explains the empirical regularity of stronger estimated quality effects as the level of aggregation of the schooling variable used in micro-earnings functions increases. This model suggests that the direction of the bias resulting from aggregating the measure of schooling quality depends on whether funds are allocated towards schools in an egalitarian or utilitarian way and on whether aggregates are constructed across or between political jurisdictions. A positive bias results if allocation is utilitarian, i.e., is aimed at maximizing the aggregate return to school quality in a political jursidiction.

Organization of the paper

This paper develops in the following way. In section I, we use economic theory to motivate particular functional forms for the log earnings equation, focusing on how quality and demand and supply shocks should be incorporated into micro empirical earnings functions. Different functional forms emerge, depending on whether human capital is measured in "efficiency units" or is assumed to be heterogeneous and depending on how quality is priced in the market. We demonstrate that standard neoclassical efficiency units models of the labor market do not rationalize the data (this point is further refined in Section IV). A richer model of the pricing of labor services is required that recognizes heterogeneous human capital and impacts of regional shocks on the pricing of the quality and quantity of human capital.

In section II, we present simple nonparametric tests to examine the basic hypothesis that quality improvements raise earnings. An inspection of the quality indices reveals that commonly used aggregate measures are not highly correlated with one another and offer conflicting evidence on the effects of quality on earnings. We test an assumption, invoked in many models in the literature, that earnings are a monotonic function of measured quality. One implication of this assumption is that within education levels, ranks of wages by states or regions of birth should coincide with ranks of quality. Using nonparametric rank tests, we test and reject this hypothesis. The results of these tests cast doubt on the validity of crude aggregated quality indices as measures of true schooling quality effects.

The pattern of rejections we find is consistent with the operation of selective migration by state of birth to states of residence. Comparisons of migrant and nonmigrant wages and tabulations of migration patterns indicate that migrants typically earn higher wages than nonmigrants and that persons from particular states of birth migrate to particular states of residence. This finding calls into question a key identifying assumption invoked in aggregate data studies: that residential location is not selective on the unobservables in destination wage equations.

In section III, we present a statistical framework for analyzing the earnings-quality relationship sufficiently broad to encompass previously-used empirical models. We apply this framework in considering the parameters of interest and sources of identifying information in several of the models developed in section I. We use studies by Card and Krueger (1992b) and Johnson and Stafford (1973) as concrete examples of different approaches to understanding the impact of schooling quality on earnings.

Section IV examines the sensitivity of estimated effects of quality on earnings to alternative specifications of the econometric model. We start with two simple models: one in which quality enters the earnings equation through intercept terms (following Johnson and Stafford (1973)) and one in which quality enters through the return to education (following Card and Krueger (1992b)). We discuss differences in functional forms, investigate the implications of various maintained assumptions, and conduct statistical tests to determine whether the data support the models under consideration. We

then develop more general models to allow for multiple channels of influence of quality on earnings, to incoporate nonlinearities in the return to education, and to account for differential effects of quality on the return to schooling in different labor markets,.

In section V, we develop a model of collective decision-making about school resources that explains why studies using aggregate quality data find stronger estimated quality effects on earnings than studies based on disaggregated measures. Our model offers an alternative to claims in the literature that measurement error in the inputs is averaged out in state aggregates, leading to more precise estimates in aggregate data studies (see Card and Krueger (1994), Betts(1995), Hanushek (1994)). This claim is unconvincing since aggregation eliminates true variation in the data and should lead to less precise estimates. In our model, maximizing behavior by political organizations that allocate educational funding gives rise to bias when aggregate quality data are used in micro earnings equations. If more funds are distributed to schools with the highest marginal returns from schooling inputs, the positive dependence engineered between the level of quality and its return will result in an upward bias for the average return to quality in micro estimates based on aggregate quality measures.

I. Introducing Schooling Quality into Earnings Equations

Most empirical studies of the earnings-education relationship are versions of the well-known Mincer log earnings model (Mincer, 1974).[1] In the Mincer model, returns to different levels of education are only implicitly determined. Persons choose their number of years of schooling by comparing earnings streams corresponding to different levels of education. Under certain conditions, this comparison leads to a linear relationship between log earnings and years of schooling,

$$\ln Y_s = \ln Y_0 + rS + U \quad ,$$

---

[1] Welch (1966) develops a unique approach. He specifies a model in which education is a factor of production, along with raw labor. Using a Lancaster-Gorman model of earnings, schooling and raw labor are two separate factors determining earnings and quality augments schooling. The coefficients on labor and quality-adjusted schooling are, respectively, the marginal product of the aggregates of total labor and the marginal product of the aggregate of schooling quality. This model implies that earnings (not log earnings) are linear functions of individual endowments of labor and adjusted schooling quality. For evidence against the Gorman-Lancaster model see Heckman and Scheinkman (1987).

where Y_s represents an individual's earnings at S years of education, and U is an unobservable. Often, additional variables such as a quadratic in post-school work experience are added to this framework.

Behrman and Birdsall (1983) suggest two ways of introducing schooling quality into this framework, which lead to the following specifications for the earnings-quality relationship:

(1) Quality entering through the rate of return to education:

$$r = r(Q) = r_o + r_1 Q + r_2 Q^2$$

so

$$\ln Y_s = \ln Y_0 + r_0 S + r_1 S Q + r_2 S Q^2 + U$$

or

(2) Quality entering through "effective" years of schooling:

$$S^* = S^* (S,Q) = w_o + w_1 S + w_2 Q$$

so

$$\ln Y_s = \ln Y_0 + r w_0 + r w_1 S + r w_2 Q + U \ .$$

In the first specification, quality enters the model only through an interaction with education, so that the effect of quality on earnings monotonically increases with the level of education. In the second model, in which quality enters independently from education, the effect of quality on log earnings is the same for all levels of education. Studies that estimate some version of these specifications include Johnson and Stafford (1973), Behrman and Birdsall (1983), and Card and Krueger (1992b).

The main shortcoming with the Mincer framework is that it does not have a clear economic interpretation as a pricing relationship in a general equilibrium context, where earnings at different education levels are determined by the demand and supply for different types of human capital and individuals can realize higher returns to their human capital through migration. (See Sattinger, 1980). Regional demand and supply forces affect earnings levels, educational attainment levels, and the rate of return to education.

Focusing only on the relationship between the rate of return to education and the quality of schooling without accounting for these other factors could lead to anomalous results. For example, schooling quality has been historically low in Southern states, but the return to education has been relatively high, presumably because of the shortage of educated labor in the region. Failure to account for labor market conditions could lead to a negative estimated relationship between educational quality and the return to schooling. States scarce in skills may have higher returns to skill as a consequence of education policies that do not favor skill formation.

Attempts to modify the Mincer framework to incorporate demand and supply shocks have been rather ad hoc, with little theoretical justification given for particular approaches. The human capital model of Ben-Porath introduces aggregate demand and supply variables into the earnings equation through the intercept of log wage equations. (See Ben-Porath (1967) and Heckman and Sedlacek (1985)). Ben-Porath models human capital in an "efficiency units" framework, in which there is assumed to be one type of human capital which individuals possess in different amounts. In section IV, we test whether the data are consistent with functional form assumptions implied from an efficiency units model and find that they are not. A richer model with heterogeneous forms of human capital is required to rationalize certain patterns of interactions found in the data.

In this section, we present some simple economic models of the pricing of schooling quality that explicitly incorporate demand and supply shocks. We take as given the quality level chosen by the state. All persons born in state b get the same quality q_b. A more complete model would consider the political economy of quality determination, a subject which we pursue in Section VI. In the text of this paper, we also assume that individual schooling levels are exogenously chosen with respect to unobservables of the earnings function. In Appendix A, we present a model of the individual choice of schooling levels for persons living in different states offering different qualities of schooling. There we also consider the implications of different quality levels on migration patterns across states.

In this section, we begin with a two skill model and then move on to consider multiple skill models. Quality is introduced into the earnings model by assuming that the amount of human capital embodied in an individual is a function of his or her educational attainment level and the quality of that education. We derive the implied functional forms for the earnings-quality relationship under alternative human capital accumulation models. The goal is to present a model with theoretical justification that is also consistent with the empirical evidence on schooling quality and earnings.

A Two-Skill Model

Assume that every individual possesses one of two skills, "0" and "1". Output in a state is assumed to be a function of aggregate quantities of the two types of skill. $\bar{S}_0$ is the aggregate of the low skill, "0", and $\bar{S}_1$ the aggregate of the high skill, "1". Assume that the technology in a state is a function of the aggregates, $\bar{S}_0$ and $\bar{S}_1$, where other factors are suppressed to focus on main ideas. Output in a given state s may thus be written as

$$Y_s = F^s(\bar{S}_{0s}, \bar{S}_{1s}),$$

and the returns to skill are given by $F_1^s = \omega_1$, and $F_0^s = \omega_0$, where the superscript denotes state s-specific variables.

Skill aggregates are defined in the following way. $\bar{S}_{0s}$ is the number of workers living in state s who have skill level "0". $\bar{S}_{1s}$ is the number of workers who have skill level "1". School quality is assumed to augment the skills of level 1 workers only. Let q_b be the average quality of skill level 1 workers educated in state b. Let N_{bs} be the number of workers educated in b who live in state s. Then $\bar{S}_{1s}$ is

$$\bar{S}_{1s} = \sum_{b=1}^{B} N_{bs} \, q_b$$

where B is the total number of birth states.

For the sake of argument, assume that an individual's state of residence is determined by a random process. A central assumption in the existing literature on measuring the impact of schooling quality on earnings is that location decisions are

random with respect to the errors in earnings equations. While there is much evidence against this assumption, which we present below, we maintain it here to understand other implicit assumptions also made in this literature.

The wages of workers born in b and living in s are

$$W_{1bs} = F_1^s q_b \qquad \text{for a "1" type worker}$$

$$W_{0bs} = F_0^s, \qquad \text{for a "0" type worker .}$$

The statistical model that corresponds to this equation is a model in which the log wage of a person i born in state b and living in state s is

$$\ln W_{ibs} = (\ln q_b + \beta_{1s}) E_i + \beta_{0s} (1 - E_i)$$

where $E_i = 1$ if the skill of person i is "1", $= 0$ otherwise and where

$$\beta_{1s} = \ln F_1^s \text{ and } \beta_{0s} = \ln F_0^s.$$

In terms of the language of the analysis of variance, this model is one in which at most two main effects, a state of birth effect and a state of residence effect, are needed to explain wages at each given skill level. At the lowest skill level no state of birth effect appears because no quality is attached to low skill workers.

Now introduce skill-neutral labor market shocks to overall productivity, λ_s, and shocks that are skill-specific, $(\lambda_{0s}, \lambda_{1s})$, so that the aggregate output equation for state s is

$$Y_s = \lambda_s F(\lambda_{0s} \overline{S}_{0s}, \lambda_{1s} \overline{S}_{1s}) .$$

In this case, the equilibrium price of skill "1" in state s is

$$\lambda_s \lambda_{1s} F_1^s = \omega_{1s} .$$

A worker born in state b residing in state s receives the wage $W_{1bs}(q_b) = q_b \lambda_s \lambda_{1s} F_1^s$. For unskilled workers, the wage is $W_{0s} = \lambda_s \lambda_{0s} F_0^s$. In logs,

$$(3) \ \ln W_{ibs}(q) = \ln \lambda_s + (\ln q_b + \ln \lambda_{1s} + \ln F_1^s) E_i + (\ln \lambda_{0s} + \ln F_0^s)(1 - E_i).$$

The corresponding statistical model is one of main effects for state of birth and state of residence <u>at each skill level</u>. There is also a common state of residence effect across all skill levels (λ_s).

In logs, the price of skill is additive in indices of aggregate variables and additive in the qualities of individual characteristics in any pricing model determined by sectoral aggregates. There is no interaction between aggregates (or regional variables) and individual characteristics, in particular quality. Demand and supply shocks are incorporated in this simple model by including indicators for the state of residence in intercepts and interacted with the skill level of education, E_i.

We test the validity of this "main effects" model using weekly wage data for three cohorts of white men from the 1980 Decennial Census (sample restrictions are described in detail in the next section and in Appendix A). Different skill levels are interpreted as different years of education. Disaggregating by education levels, we test for the main effects representation by adding interactions between regions of birth and regions of residence to the model and testing whether the associated coefficients are jointly significantly different from 0. We also test whether the state of birth main effects are equal across states of birth and whether the state of residence main effects are equal across states of residence. Table 1(a) presents the table of interactions. Table 1(b) contains the F-statistics and associated p-values for testing the hypothesis of no interaction, which indicate that each of these restrictions is strongly rejected. In section IV, we test and reject another implication of this model--that estimated quality effects are the same across all states of residence. There is much evidence supporting the notion of differential effects of quality on log wages in different regions of residence - contrary to the model of equation (3).

We also find evidence of a state of residence effect on the returns to education (the coefficient $\ln F_1'$ on E_i). Disaggregating by education levels, we find little evidence of state-of-residence effects for highly educated workers (those with more than one year of college) but substantial state-of residence effects for less educated workers (those workers with 12 or less years of schooling).

Efficiency Units Models

The widely-used efficiency units specification assumes that output in a state depends on the state aggregate stock of human capital, $\overline{H}_s$. There is one type of human capital that can be aggregated across individuals who possess it in different amounts. The return to a unit of human capital in state s (allowing for shocks to arguments of skill in the production function) is

$$R_s = \lambda_s F_s{}'(\lambda_s \overline{H}_s)_.$$

A person's own stock of human capital, H_{ib}, depends on the quality invested in the person and on his or her years of education. One conventional assumption is that a constant amount of human capital is accumulated per year of schooling. A person born in state b with E_i units of schooling gets a total of $q_b E_i$ units of quality (assuming the same quality for each year of education and no depreciation). The function defining human capital and the implied wage equation are, respectively,

$$H_{ib} = q_b E_i \ ,$$

and

$$W_{ibs} = R_s H_{ib} = R_s q_b E_i \ .$$

In logs, there is no interaction between quality and education. In a log wage equation, the effect of schooling quality operates through the main effects for state of birth, while demand and supply factors operate through main effects for state of residence:

$$(4) \quad \ln W_{ibs}(q) = [\ln \lambda_s + \ln F'^s(\lambda_{1s}\overline{H})] + \ln q_b + \ln E_i \ .$$

Thus, a multiplicative transformation function for human capital implies a functional form for the earnings equation in which schooling quality operates through the state of birth intercepts ($\ln q_b$) and not through the rate of return to education (the coefficient on E_i). An implication of the model is that the rate of return to education is the same across all states of birth and states of residence. Demand and supply can be captured in log wage model by including a state of residence fixed effect (the term in brackets).

Alternative assumptions about the form of the function translating schooling inputs into human capital lead to alternative implied channels of influence for the impact of

quality on earnings. One such function assumes that quality is "self-productive." A person with E_i years of schooling has human capital

$$(5) \quad H_{ib} = (1 + \eta \, q_b)^{E_i}, \quad \eta > 0.$$

Thus, for a given quality level, human capital accumulates exponentially with additional years of education. This specification captures the notion that quality augments the rate of growth of knowledge obtained from schooling.

The log wage equation corresponding to this specification introduces quality as a determinant of the rate of return to education. Wages for person i born in state b and living in state s are

$$W_{ibs} = R_s H_{ib} = R_s (1 + \eta \, q_b)^{E_i}$$

so in logs

$$\ln W_{ibs} = \ln R_s + E_i \ln(1 + \eta \, q_b)$$

$$(6) \quad \ln W_{ibs} \approx \ln R_s + \eta q_b E_i.$$

As with the multiplicative human capital accumulation model, state-wide demand and supply shocks are again in the term $\ln R_s$, which corresponds to a state of residence fixed effect. In logs, there is no interaction between the state of residence effect and education--contrary to the empirical evidence, presented in section IV, which reveals substantial patterns of interaction. However, the quality variable now enters the log wage equation interactively. State differences in the transformation function for human capital could be incorporated through a state-specific parameter $\eta_{b'}$. Factors that could lead to different values of η_b across states include differences in family and community environments.

Aggregative Heterogeneous Human Capital Models

In the models considered thus far, demand and supply effects on log wage equations operate only through state of residence intercepts. Relaxing the efficiency units assumption alters the empirical specification of wage equations in an important way. We now derive the functional form for the log wage equation implied by a model with different types of human capital receiving different prices in equilibrium. This model

rationalizes the patterns of interaction found in Table 1. We assume a geometric human capital transformation function, given by (5), for each type of human capital.

The following aggregate production technology combines features of the skill accumulation model discussed previously, and a heterogeneous human capital model:

$$Y_s = \lambda_s F(\lambda_{1s} \overline{H}_{1s}, ..., \lambda_{Js} \overline{H}_{Js})$$

where $\overline{H}_{\ell s}$ is the aggregate stock of human capital of type ℓ in state s. Let $E_{i\ell}$ be years of education for person i accumulating human capital of type ℓ. Personal accumulation follows the growth process for person i born in state b:

$$H_{i\ell b} = (1 + \eta_\ell q_b)^{E_i} \quad .$$

In this case, the transformation function for human capital differs by type of human capital but is the same across states of birth ($\eta_{\ell b} = \eta_\ell$). We assume that individuals specialize in one type of human capital, and that different values for η_ℓ and differential costs of investment lead to a different number of years invested for different types of human capital. For example, an individual observed with 16 years of education is assumed to possess a different type of human capital than someone with 12 years of schooling, but within each education level, there may be important differences in quality.

Marginal productivity pricing produces the following specification of wages for workers of type l:

$$W_{ib\ell s} = \lambda_s \lambda_{\ell s} F'_\ell H_{i\ell b}$$

so the implied wage equation is:

$$(7) \qquad \log W_{ib\ell s} \cong \ln \lambda_s + \ln \lambda_{\ell s} + \ln F'_\ell + \eta_\ell q_b E_{i\ell} \ .$$

In this framework, $\ln \lambda_s$ is a common state of residence effect, $\ln \lambda_{\ell s} + \ln F'_\ell$ is a state of residence - level of education interaction, and $(\eta_\ell q_b)$ is a level-of-education-specific return to education that depends linearly on quality. Note that there is no state of residence-quality interaction in this model, i.e., there are no state of birth-state of residence interactions in logs.

In this model, log wages would be linear in the years of education if $\eta_\ell = \eta$ for all levels of education. Exact linearity between levels of education and state of residence

would be produced if restrictions on aggregate shocks and technology such as the following were satisfied:

$$\ln \lambda_{ts} + \ln F_t^s(s) = \chi_{s0} + \chi_{s1} E_{it} \quad .$$

In this special case,

$$\ln W_{it_{bs}} = \ln \lambda_s + \chi_{s0} + (\chi_{s1} + \eta) E_{it} \quad .$$

Most of the models estimated in the school quality literature assume that log earnings is linear in education.[2] While linearity can be justified under the strict conditions of the older form of Mincer model (Mincer 1958), it is not easy to rationalize it in a general equilibrium setting when demand and supply shocks are introduced. As will be shown in section IV, the empirical evidence speaks against linearity in all three Census years, so that a more general specification appears to be the appropriate one. The departure from linearity may be either in the personal accumulation process ($\eta_t \neq \eta$) or in the way markets price different education levels. There is little justification for the assumption that F_t^s is linear in years of education. Such linearity would require strong assumptions on the aggregate production technology.

This model, like the previous one, does not rationalize the evidence reported in Tables 1 (a) and (b) demonstrating a statistically significant region of birth-region of residence interaction for log wage equations. Below we document that part of this interaction - but by no means all - is due to differential effects of quality on the return to education in different regions. Persons born in the same region have returns to their education that depend on their region of residence. Such evidence speaks strongly against any linear aggregate of quality either within or across educational levels such as are widely used in aggregative analysis.

There are at least two reasons why such interactions might arise.

(1) If state of birth proxies quality of schooling, a hedonic model specific to each region, educational level and quality level would be consistent with the finding of

---

[2] An exception is a study by Wachtel (1976), in which the quality relationship is estimated separately by educational attainment level.

significant interactions. Different bundles of education and quality would be priced out differently across state and regional labor markets.

(2) Below we present evidence that quality has different effects on the return to schooling in different regions. If there is comparative advantage in different regions of residence for persons born in different regions, say because of costs of migration or because of the specificity of community and human capital and its suitability in alternative regions, the observed region of birth - region of residence patterns in the earnings equation could be produced. Below we also present evidence that distance effects explain part of the observed interaction. Such effects arise from the operation of selective migration. Persons who receive good wage offers in distant regions can afford the direct and psychological costs of traveling long distances to take a job. In section IV, we demonstrate that conventional estimation strategies in the literature ignore this interaction and other region of birth-region of residence interactions. When they are taken into account, estimated effects of schooling quality on earnings are affected, sometimes dramatically so.

II. Testing Basic Premises of the Quality Hypothesis

Data Description

Our observations on individuals' earnings and demographic characteristics come from the 1970, 1980, and 1990 micro Census samples. The measure of earnings is the weekly wage. To facilitate comparisons between of our empirical results and those of Card and Krueger (1992b), who analyze ten-year birth cohorts of white men from the 1980 Census, we impose identical sample restrictions. Appendix B summarizes the sample exclusion restrictions and gives exact sample sizes for the different extractions (which range from around 100,000 to nearly 600,000). Our analysis samples consist of white men born in the U.S. in the years 1910-1959 who worked at least one week and reported positive wage and salary income within a certain range. This time frame covers five birth cohorts: 1910-1919, 1920-1929, 1930-1939, 1940-1949, and 1950-1959.

Our measures of adult earnings are linked to schooling quality measures by relating individual level data on weekly wages to state level average measures of schooling inputs for the years in which the individual would have attended school (grades K-12). It is assumed that individuals are educated in their state of birth. For example, a person born in Florida in the 1920-29 cohort is assigned an index of that states' schooling inputs averaged over the years in which his birth cohort attended school (1926-1947). We consider four empirical measures of quality in this study: two measures of the pupil-teacher ratio (one based on pupil enrollment and one based on average daily attendance), term length, and relative teacher salary. For the earlier cohorts, who were not subject to mandatory schooling laws, and for more rural areas with family farming demands, the attendance-based pupil-teacher ratio is a better proxy for true quality.[3] All of quality measures are derived from published biennial reports by the U.S. Department of Education.[4]

Nonparametric Tests of The Schooling Quality Hypothesis

The premise of the quality-earnings hypothesis is that higher quality schooling implies higher earnings. This premise alone is almost a tautology, because "quality" per se is something that yields a positive benefit to the recipient - in this case higher earnings. To make the premise empirically testable, we must further assume that quality levels can be measured in terms of such indices as pupil-teacher ratios, teacher salary, and term length. This form of the hypothesis can be tested, since the variables in question are observable.

Before testing the hypothesis, however, it is instructive to look at the aggregate measures of schooling quality used in the literature. Unfortunately, the frequently used indices are sometimes inversely related to one another. This may naturally occur if there

---

[3] Most state mandatory school attendance laws were not passed until the early 1920s. Sometimes state or district educational funding depended upon enrollment, in which case there was an incentive to inflate enrollment numbers.

[4] There are some substantial differences between the quality measures we use and those used by Card and Krueger (1992b), although the data sources are the same. These are discussed in detail in Appendix C. Briefly, we focus on the attendance based pupil-teacher ratio while they consider only the enrollment based one. Additionally, our term length measure is race specific for the southern segregated states, and we use a different deflator for the relative teacher salary measure.

are budget constraints on schooling authorities. As teacher salaries rise, there may be a concurrent increase in classroom size as school boards attempt to maintain their budgets. That is, one measure may indicate an improvement in quality while another indicates a decline. Table 2, which tabulates rank correlations in quality measures across states, shows that alternative measures do not move together. Nevertheless, since these are the best aggregate measures available, we assume, as does the literature, that they capture some aspects of "true" quality.

If quality has a uniform effect on earnings at all education levels, as in the multiplicative-efficiency-units earnings model given in equation (3), the wages of two persons with the same years of education, living in the same labor market but from different states of birth, should be ranked according to the quality rankings of their states of birth. If schooling quality affects only intercepts, as in the geometric-efficiency-units model in (4), or only slopes as in the heterogenous model in (7), the wage rankings should also coincide with quality rankings.[5]

Assuming that individuals are educated in their state of birth, the hypothesis that wage ranks and quality ranks of birth states should be correlated can be tested nonparametrically. Table 3 presents Kendall-Tau correlations (and p-values in parentheses) between the rankings of four different state of birth quality measures and the rankings of mean wages by states of birth for individuals with 16 years of education or more. Correlations are calculated separately within regions of residence to control for regional labor market effects on earnings.[6] The frequent sign reversals displayed in the tables suggest no relationship between the quality and mean wage rankings. Using other education levels and medians instead of means yields similar results. The lack of any correlation calls into question our ability to measure levels of schooling quality with the crude aggregate indices widely used in the literature.

---

[5] One possible case in which the wage and quality rankings may not coincide is when the earnings model includes fixed effects for state of birth as well as slope effects which depend on quality. This type of specification is adopted by Card and Krueger (1992b). The rankings of wages may not correspond to the quality ranks in this example because they could be undone by the intercept effects. However, even in this case, the relative ranks of wages of persons classified by states of birth should be identical across all labor markets.

[6] We take the groups of states classified into nine divisions by the Census to be regional labor markets.

Suppose school quality determines a person's wage ranking, but that it cannot be measured by the standard aggregate indices. Another testable implication of the quality-earnings hypothesis, one that avoids the problem of measuring quality, is to ask whether persons from a particular state of birth tend to have the same relative rank in the wage distributions across all regions of residence. The quality-earnings hypothesis would predict invariance in the state of birth wage rankings across all regions of residence.[7] Table 4 presents Kendall Coefficient of Concordance statistics (and p-values), that are used to test for stability in the region of birth wage rankings across regions of residence. They provide some support for this implication of the hypothesis; the 1990 Census data show several significant p-values. Thus, while the rankings of empirical measures of schooling quality and earnings are uncorrelated, rankings of earnings by region of birth give some evidence of stability across regions. Of course this evidence is consistent with the operation of other state of birth-specific factors such as family background effects or other state of birth-specific effects.

It is useful to ask what explanations are consistent with the evidence from the rank tests -- particularly the finding of no correlation between quality and wage rankings. One possible explanation is that states differ in their production technologies for transforming schooling inputs (the variables that we measure) into human capital. This notion is consistent with Hanushek's (1986) observation that efficiency in translating educational inputs into outputs, and not just the absolute levels of inputs, is important. Factors that could lead to varying levels of efficiency across states include the strength of teacher unions, state laws governing education, and teacher proficiency. Different education-quality production functions could explain the observed nonmonotonicity of earnings in measured quality inputs.

The results of the rank tests can also be explained if individuals migrate selectively from states of birth to particular states of residence, based on prospective gains and individual-specific costs. Suppose persons born in different regions receive different quality endowments, face different costs of migrating, and expect different payoffs in

---

[7] When making wage comparisons across different labor markets, rank comparisons are much more robust than level comparisons since they are not sensitive to regional differences in the cost of living.

destination regions. Even if higher schooling quality raises productivity in all regions, the pattern of migration choices need not imply strict dominance of the wages of persons from high quality origins in all destination regions. Persons migrating long distances will tend to have higher earnings in all destination regions than those persons making short moves or no moves at all. Below, we find empirical evidence that the wage data are consistent with patterns produced by selective migration.

Nonmonotonic school quality-earnings relationships, the operation of selective migration, or measurement error in the quality indices could account for the lack of correlation between wage and quality rankings. This evidence casts doubt on the existence of invariant aggregated schooling quality measures that successfully predict relative achievement in different economic environments. We next present more conclusive evidence that migration is selective on wages and discuss the implications of this evidence for empirical procedures used to estimate the effect of quality on earnings.

Selective Migration

In aggregate data studies, interstate migrants play a major role in separating out the effect of quality on earnings from the effects of other factors. Since aggregating the quality data to the state level eliminates intrastate variation in schooling quality, migration across states is needed to distinguish state of birth quality effects from state of residence labor market effects. (See the discussion in Heckman, Layne-Farrar and Todd, 1995). When individuals migrate randomly, then the variation in quality within a given state of residence is exogenous and a quality effect on earnings can be identified. This is the most favorable case for aggregate estimation strategies. But a more realistic view is that migrants consider the demand for their skills in the destination region. In response to an increase in the demand for human capital (skilled labor), there could be migration from states with higher schooling quality to the states demanding such skills. If selective migration is important, then the effects of schooling quality can only be isolated using empirical procedures that control for selection. In Appendix A, we present a model of migration across regions that closes the two skill model of Section I. The model predicts migration from high quality of schooling states to low quality states.

If migration were random, there should be no systematic difference between the annual wages of migrants and nonmigrants; yet we find an enormous difference. Table 5 presents the p-values from Wilcoxon pairwise nonparametric rank tests, used to test whether the wage distributions of migrants and nonmigrants are equal. The tests, performed separately within labor market regions and within educational classes, provide strong support for selective migration. The migrant wage distribution stochastically dominates that of nonmigrants for every cohort, educational level and almost every regional labor market (there are some cases where we do not reject for the Pacific region).

Table 6, which tabulates the migration patterns for the 1980 Census samples, shows that persons born in a region are much more likely to live in it. Those who do migrate are likely to go to neighboring regions, so distance between regions plays a role in the migration decision. This relationship is stronger for persons with less education and for persons born in the Pacific region. Other Census years reveal the same pattern. In work not presented here, we also find some evidence that the quality levels of migrants are higher than the quality levels of nonmigrants -- especially at lower schooling levels.[8] Enhanced mobility is one benefit of receiving higher quality schooling, but it is a liability for economists seeking to identify schooling quality effects using aggregate measures of quality in micro earnings functions.

III. A Statistical Framework for Organizing the Evidence

Section I motivated particular functional forms for the earnings-quality model with economic theory. In this section, we develop a statistical framework general enough to encompass all of the models introduced in Section I and most of the other models used in the literature.

Express the logarithm of earnings of individual i born in state b, in birth cohort c, residing in state s as

$$(8) \quad y_{isbc} = \tilde{\theta}_{isbc} + X_{isbc}\tilde{\beta}_{sbc} + E_{isbc}\tilde{\alpha}_{sbc} + \varepsilon_{isbc}$$

---

[8]We compute the average quality levels of migrants as compared to that of nonmigrants. We find that in the majority of cases, the mean quality of migrants exceeds that of natives. This pattern also holds when median values are used. See Heckman, Layne-Farrar, and Todd (1995) for complete results.

$$i = 1,...,I; \quad s = 1,...,S; \quad b = 1,...,B, \quad c = 1,...,C$$

where $\tilde{\theta}_{sbc}$ is an intercept term; E_{isbc} is education --which may be a vector or a scalar; $\tilde{\alpha}_{sbc}$ is the associated coefficient; and X_{isbc} represents other regressors such as work experience with associated coefficient $\tilde{\beta}_{sbc}$. It is conventional to assume that E is a scalar determinant of earnings - education - but our analysis can readily be extended to the vector case. For simplicity of exposition, in this section we treat E as a scalar variable, but in our empirical analysis we consider both the vector and scalar cases. ε_{isbc} is a mean zero variable unobserved by the analyst that may or may not be known to person i. Associated with s, b, and c are vectors of dummy variables, D_{si}, D_{ci}, and D_{bi}, that indicate the appropriate s, b and c subscript for person i.

As a statistical model of mean earnings, equation (8) is quite general. If E_{isbc} were a vector of dummy variables indicating different levels of educational attainment, (8) would impose no restrictions on mean earnings at all. Each education - state of residence - state of birth - cohort of birth cell would be fully described by (8). Linearity in E_{isbc} imposes a stringent restriction which we show below is strongly rejected by the data.

Both the intercept and the slopes can be decomposed into an analysis of variance structure:

$$(9) \quad \tilde{\theta}_{sbc} = \theta_c + \theta_b + \theta_s + \theta_{bc} + \theta_{sc} + \theta_{sb} + \theta_{sbc}$$

$$(10) \quad \tilde{\alpha}_{sbc} = \alpha_c + \alpha_b + \alpha_s + \alpha_{bc} + \alpha_{sc} + \alpha_{sb} + \alpha_{sbc}$$

$$(11) \quad \tilde{\beta}_{sbc} = \beta_c + \beta_b + \beta_s + \beta_{bc} + \beta_{sc} + \beta_{sb} + \beta_{sbc}.$$

Persons may also have idiosyncratic components of θ, α and β reflecting differences at the individual level with regard to birth cohort, state of birth, and state of residence effects and their interactions. For example, persons may differ in their unmeasured quality of schooling, which is known to vary widely within birth cohorts.

Estimated coefficients are **population average** effects only if the indicator variables are uncorrelated with the deviations of persons around group means. Problematic in this regard are state of residence variables (dummy variables denoted D_{si}) and the associated

interaction variables. If a person's choice of residence is based at least in part on their ε-deviations from average outcomes, least squares is inconsistent for the population mean and estimates instead the mean among those who chose to migrate/stay in states. We have already shown that selective migration appears to be an important phenomenon in the data, and we present futher evidence on this question below.

In our empirical work, we adopt conventional normalizations in the analysis of variance and we measure effects around omitted categories. The most general analysis of variance form of model (6) is fully saturated, so there are no remaining degrees of freedom to estimate a quality effect, since quality is identical to cohort and state of birth. Empirical studies differ in the types of restrictions they impose to produce estimates of a quality effect.

The precise nature of the restrictions required to identify a quality effect depends on how quality is assumed to affect earnings. If schooling quality operates independently of the level of education, such as in the earnings model derived from the multiplicative human capital accumulation model (given in (3)), then it operates through θ. Studies of schooling quality relying on aggregate measures use quality defined at aggregate geographic levels and assume that a common quality component applies to all persons born in state b in cohort c. "θ_{bc}" would be replaced by some lower dimensional function of measured quality, Q_{bc}, e.g.,

$$(12) \quad \theta_{bc} = Q_{bc} \, \delta \, .$$

More general functions of Q_{bc} could be estimated provided that they do not exhaust the $(B - 1)(C - 1)$ degrees of freedom in θ_{bc}. . Observe that this specification would imply that ranks of wages by state of birth and ranks of quality measures by state of birth should be highly correlated for each education level in each labor market, contrary to the evidence presented above.

Schooling quality could also operate through the slope coefficients. This type of earnings specification is implied by the heterogeneous human capital model, equation (7). A discussion parallel to that presented for θ then applies to α.

Restrictions have to be placed on α_{bc} to "make room" for Q_{bc} in the model. A prototypical restriction is the one used by Card and Krueger (1992b):

$$(13) \quad \alpha_{bc} = Q_{bc}\, \varphi$$

Again, more general models could be fit as long as the $(B-1)(C-1)$ degrees of freedom in α_{bc} are not exhausted. The lack of any association between ranks of wages by state of birth and ranks of quality by state of birth within education levels and within labor markets, which was established in Section II, argues against this specification.

Third, schooling quality could operate on earnings both through θ and α. An advantage of focusing on α, stressed by Card and Krueger, is that it permits application of a classical "fixed effects" estimation strategy. One can control for θ_{bc} while estimating α_{bc}.[9] By observing persons with the same α_{bc} working in different state labor markets s, we can identify a quality effect free of conventional fixed-effects bias, assumed to operate through intercepts and not slopes. By estimating models with state-of-birth-specific fixed effects in the slope coefficients, as Card and Krueger do in their analysis, it is also possible to control for fixed effects in slopes as well.

A crucial implicit assumption that is invoked to secure identification using the fixed effects assumption is that the dummy variables indicating state of residence are uncorrelated (mean independent) with the error term. This rules out selective migration, i.e. it assumes that state of residence dummies (D_{si}) are mean independent of ε_{ibsc}.

$$E(\varepsilon_{ibsc} \mid D_{si}) = 0 \, .$$

Fourth, schooling quality might enter the earnings model through other channels. The average quality of persons in a labor market could affect earnings through an aggregate supply effect on factor prices. In this case, the state of residence components (both the levels (θ_s, θ_{sc}) and slopes (α_s, α_{sc})) could be functions of aggregate quality. Since most studies focus only on one channel of influence, they arguably understate or overstate the contribution of quality to earnings, depending on whether quality increases or decreases the neglected components of the pricing equation. We later present evidence for certain models that increasing the aggregate supplies of quality in a state *depresses* marginal prices of skill in the state, at least for certain dimensions of measured quality.

---

[9] This is also true in a more general model in which E is a vector of dummy variables. In this case, the θ components would be the α components for a benchmark state.

Aggregate schooling quality is just one of several aggregate economic variables that are plausible determinants of the state of residence components (s-subscripted effects). In section IV, we present new evidence on the impact of aggregate stocks of human capital, physical capital, industrial structure and schooling quality on state-wide and regional earnings equations. Aggregate economic variables play an important role in determining pricing equations. This portion of our analysis continues a tradition of incorporating aggregate demand variables into micro earnings equations begun by Tinbergen (1975, 1977) and Sattinger (1980).

Two Important Special Cases

We now discuss the restrictions used in two representative studies for introducing schooling quality into an earnings equation. The first study we consider is that of Johnson and Stafford (1973) who analyze the effect of quality operating through intercepts. The second study we consider is the one by Card and Krueger (1992b), who introduce quality into the earnings function through slope coefficients. One restriction that is imposed in both models is that the coefficients on demographic variables other than education are constant across states of birth and states of residence, but are allowed to vary across cohorts ($\beta_{sbc}=\beta_c$). This restriction is tested and rejected in Section IV. Another common feature of these two models, like others estimated in the literature, is that log earnings are assumed to be linear in education. We test and reject this assumption. We find that removing the restriction of a linear relationship between earnings and education has important consequences for the estimates of the impact of quality on earnings.

The Johnson-Stafford Model

We present the Johnson-Stafford (JS) model as a special case of equation (8), first considering their restrictions on θ and then on α. In the JS model, quality increases log earnings independently of the level of education, so that quality operates through the $\tilde{\theta}_{sbc}$

components.[10] This linearly additive specification for quality can be motivated within the multiplicative human capital accumulation model described in section I (see equation (3)). The JS model assumes first that there are no state of residence effects in levels and no state of birth-state of residence interactions:

$$\text{JS-1: } \theta_{sb} = 0, \ \theta_{sc} = 0, \ \theta_{sbc} = 0.$$

To simplify notation, define $\tilde{\theta}_{bc} = \theta_c + \theta_b + \theta_{bc}$. Johnson and Stafford identify a quality effect by assuming second that quality operates through the state of birth intercept components:

$$JS - 2: \ \tilde{\theta}_{bc} = \theta_c + \varphi Q_{bc}.$$

Finally, they assume that the marginal return to education is cohort-specific but does not depend on region of residence or state of birth:

$$\text{JS-3: } \alpha_{sbc} = \alpha_c.$$

Assumptions JS-1 through JS-3 are testable. Johnson and Stafford test and do not reject JS-3 with their data. They do not test JS-1. Using Census data we test and reject both JS-1 and JS-3. There are strong state of birth - region of residence level and interaction effects in the price of schooling term. For the sake of brevity we do not report these results here.

The Card and Krueger Model

The Card and Krueger (CK) model places a fundamentally different interpretation on the intercept parameters than does the JS model. State of birth equation intercepts are interpreted as "nuisance parameters" included to account for possible "fixed effects" bias. Quality is assumed to affect earnings only through the state of birth components of the slope parameter on education, α_b and α_{bc}. However, this distinction is arbitrary. Both α and θ may be a source of information about quality effects on earnings or a source of bias. Indeed, α could be a nuisance parameter for θ. Therefore, when we later estimate a

---

[10] Johnson and Stafford (1973) consider adding an interaction between schooling quality and education and find it to be insignificant. We therefore present the restrictions inherent in their preferred specification (given by column 1b in Table 2 of their paper).

26

version of the Card and Krueger model, we consider the effects of quality both on intercepts and slopes.

Restrictions on θ

The Card-Krueger (CK) model is also developed as a special case of equation (8). The CK model places only two sets of restriction on θ. First, they assume that there are no interactions between state of birth and state of residence:

$$\text{CK - 1: } \theta_{sb} = 0, \theta_{sbc} = 0.$$

Thus, they abstract from the possibility that persons born in state b have a comparative advantage working in certain states. Lower migration costs could be one source of comparative advantage.

Restrictions on α

Second, parallel to their restriction on θ, they assume that

$$\text{CK-2: } \alpha_{sb} = 0 ,$$

so that persons born in b do not have a comparative advantage or disadvantage in s.

To reduce the dimension of their empirical problem they further assume that state of residence effects are identical within regions:

$$\text{CK - 3: } \alpha_{sb} = \alpha_{r(s)b}, \alpha_{sbc} = \alpha_{r(s)bc} \text{ for all s in r}$$

where $r(s)$ denotes a region (a collection of states).

Let the state of birth component of the education return, inclusive of the cohort mean, be given by $\tilde{\alpha}_{bc} = \alpha_b + \alpha_c + \alpha_{bc}$. The quality effect is defined by restricting α_{bc} to satisfy equation (13). The key restriction is that the (B-1)(C-1) terms in the matrix of interactions lie on a line. This is not at all restrictive if $B = C = 2$, but becomes very restrictive when (B - 1) (C- 1) is high dimensional. This restriction leads to

$$\text{CK-4: } \tilde{\alpha}_{bc} = \alpha_b + \alpha_c + \varphi Q_{bc} .$$

Observe that a weaker assumption would subscript φ:

$$\text{CK-4': } \tilde{\alpha}_{bc} = \alpha_b + \alpha_c + \varphi_c Q_{bc} ,$$

so φ is cohort-specific. Similarly, one could make φ birth state-specific. In some versions of the model the α_b are suppressed so that the estimated effect of quality operates through

main effects (α_b) and interactions (α_{bc}). Finally, for the sake of completeness we record their implicit exogeneity assumption:

$$\text{CK-5: } E(\varepsilon_{ibsc} \mid D_{si}, D_{bi}, D_{ci}) = 0,$$

where D_{si}, D_{ci} and D_{bi} are, respectively, state of residence, cohort and state of birth dummies for individual i.

The Card and Krueger Parameterization

The key assumptions invoked by the CK model have just been discussed. Their estimation procedure, however, implies additional implicit assumptions not intrinsic to the specification. These assumptions arise from the use of a two-stage estimation procedure.

In their first stage, they run regressions of y on X and E and decompose $\tilde{\alpha}_{r(s)bc}$ into within-region and between-region components, but they do not impose assumption CK-4. That restriction is imposed in the second stage. However, they impose restrictions CK-1, CK-2 and CK-3 on the first stage to obtain a <u>restricted</u> analysis of variance model

$$\tilde{\alpha}_{r(s)bc} = \tilde{\alpha}_{r(s)c} + \tilde{\alpha}_{bc}$$

where for simplicity we define $\tilde{\alpha}_{r(s)c} = \alpha_{r(s)} + \alpha_{r(s)c}$ and $\tilde{\alpha}_{bc} = \alpha_c + \alpha_b + \alpha_{bc}$.

Letting $w(r \mid c)$ be the proportion of cohort c living in region r, Card and Krueger impose the normalization

$$\text{CK-6: } \sum_{r=1}^{R} \tilde{\alpha}_{r(s)bc} \, w(r|c) = 0$$

which is equivalent to running regressions of the estimated coefficients on regional dummy variables for each cohort. This removes the regional mean from the estimated rate of return. Thus, $\tilde{\alpha}_{bc}$ is equal to the rate of return deviated from the appropriate regional mean. As discussed in the next section, deviations are taken around a regional mean to control for demand-driven regional labor market shocks on the rate of return, although such deviations also remove aggregate supply shocks that arise from differences among regions in schooling policies.

In the second stage regression, Card and Krueger regress estimated values of $\tilde{\alpha}_{bc}$ on dummies for state of birth, cohort of birth and measures of schooling quality,
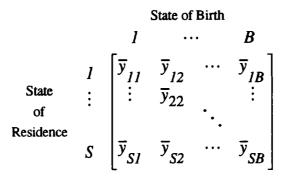
$$\tilde{\alpha}_{bc} = \alpha_b + \alpha_c + \varphi \, Q_{bc} \, .$$

In the second stage, however, they fail to deviate the regressors from the regional mean. This can be shown to bias downward in absolute value the estimated effects of quality (Q_{bc}) on the rate of return measured from the region mean $(\tilde{\alpha}_{bc})$, but our empirical analysis suggests that the effects of this error are minor in any event. If national and regional means coincide, there is no bias.[11]

Understanding The Sources of Identification

A simple table clarifies the identification strategy used by Card and Krueger and our evidence against it. For simplicity consider only a single cohort and for a given level of education, say 12 years. Let $\bar{y}_{sb}$ be the mean log earnings for people born in b and living in s. Imagine a S × B table of means as depicted in Table A.

Table A

State of Birth

$$
\begin{array}{c}
 & \quad 1 \qquad \cdots \qquad B \\
\begin{array}{c} 1 \\ \text{State} \\ \text{of} \\ \text{Residence} \\ S \end{array}
\begin{bmatrix}
\bar{y}_{11} & \bar{y}_{12} & \cdots & \bar{y}_{1B} \\
\vdots & \bar{y}_{22} & & \vdots \\
 & & \ddots & \\
\bar{y}_{S1} & \bar{y}_{S2} & \cdots & \bar{y}_{SB}
\end{bmatrix}
\end{array}
$$

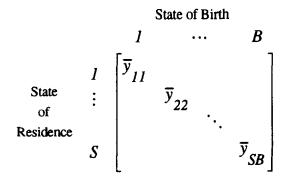If all people who are born in b remain in b their entire lifetime the table tooks like this:

---

[11] Speakman and Welch (1994) make the correct point that, in general, it is necessary to deviate the second stage regressors around regional intercepts to obtain consistent estimates of φ. Our empirical evidence shows that the impact of not deviating the regressors is minor, but we do so in all of our analyses.

State of Birth

$$
\begin{array}{ccc}
1 & \cdots & B
\end{array}
$$

$$
\text{State of Residence} \quad
\begin{array}{c} 1 \\ \vdots \\ S \end{array}
\begin{bmatrix}
\bar{y}_{11} & & & \\
 & \bar{y}_{22} & & \\
 & & \ddots & \\
 & & & \bar{y}_{SB}
\end{bmatrix}
$$

All the information in the table is on the diagonal. All of the off-diagonal terms are missing.

Assuming an additive model, mean earnings can be written as,

$$\bar{y}_{sb} = \mu + \alpha_b + \beta_s$$

where α_b is a state of birth effect and β_s is a state of residence effect. This model assumes that there is a state of birth effect common to persons from different states of residence and a state of residence effect common to persons from different states of birth. If no one migrates, then all b = s and we can never separate α_b from β_s, ie. , state economy - specific factors would be confounded with state-specific quality factors such as parental background or pupil-teacher ratios.

To identify a state of birth endowment effect, α_b, which is assumed to be generated by schooling quality, it is necessary to have information on the off diagonal terms. If there are two or more state of birth cohorts residing in a state of residence s, we can eliminate the common state of residence effect to identify α_b relative to $\alpha_{b'}$ for two states of birth b and b′:

$$\bar{y}_{b,s} = \mu + \alpha_b + \beta_s$$
$$\bar{y}_{b',s} = \mu + \alpha_{b'} + \beta_s$$

so

$$\bar{y}_{b,s} - \bar{y}_{b',s} = \alpha_b - \alpha_{b'}.$$

One could identify the change in the value of α_b from one state of birth to another using $\alpha_b - \alpha_{b'}$ so estimated. By taking differences within means indexed by a common b, it is possible to identify β_s.[12] In some specifications considered by Card and Krueger, this is what is done, and α_b is parameterized as a function of quality in the state of birth Q_b. When there is information on more than one cohort, one can estimate α_b for each cohort, $\tilde{\alpha}_{bc}$, where it should be recalled that "~" denotes the mean returns taken with (b,c) cells before before main effects are eliminated. Taking deviations of $\tilde{\alpha}_{bc}$ around cohort and state of birth means gives,

$$\tilde{\alpha}_{bc} - \alpha_b - \alpha_c = \alpha_{bc}$$

where α_{bc} was implicitly defined earlier in equation (10). Card and Krueger assume that this interaction is a linear function of the quality of education for each cohort (c) (Q_{bc}) so

$$(*) \qquad\qquad \alpha_{bc} = \varphi\, Q_{bc}\ .$$

φ is identified by comparing two or more states of birth, i.e.,

$$(**) \qquad\qquad \alpha_{bc} - \alpha_{b'c} = \varphi(Q_{bc} - Q_{b'c}).$$

Given the strict functional form, φ can be identified within cohorts, or across cohorts. Thus it is possible to extend this model to estimate a specific φ or a b-specific φ by going across cohorts or by comparing any two state of birth cohorts e.g.

$$\alpha_{bc} - \alpha_{bc'} = \varphi(Q_{bc} - Q_{bc'})$$

or

$$\alpha_{bc} - \alpha_{b'c''} = \varphi(Q_{bc} - Q_{b'c''}).$$

By dividing the left hand side by the quality contrast, φ can be identified. In fact, it is clear that if the contrasts are taken only within cohorts, a c-specific φ, φ_c, can be identified. Similarly, if contrasts are taken within b groups, a b-specific φ, φ_b, can be identified.

Crucial to their strategy is the assumption that there are no interactions between state of birth and state of residence, i.e. that in a model with interaction

$$\bar{y}_{bs} = \mu + \alpha_b + \beta_s + \Delta_{bs}\ ,$$

---

[12] Assuming the normalizations $\sum_b \alpha_b = 0$ and $\sum_s \beta_s = 0$ identifies each component relative to μ. Otherwise, the components are identified relative to an omitted category.

$\Delta_{bs} = 0$. If this is not true, then

$$\bar{y}_{b,s} - \bar{y}_{b',s} = \alpha_b - \alpha_{b'} + \Delta_{bs} - \Delta_{b's} \ .$$

Estimated relative state of birth effects now depend on choices of the states of residence. Although unlikely, it might just happen that model (**) is still true if unit increases in quality have the same effect on the rate of return across all states of residence. That is a testable restriction which we reject in 1970, 1980 and 1990. More generally, we test for the absence of interaction ($\Delta_{bs} = 0$) and reject it in all years we consider. The presence of such an interaction is consistent with patterns of comparative advantage for individuals from certain states of birth in certain states of residence, inducing individuals to migrate to exploit favorable economic opportunities.[13] Allowing for non-zero interaction ($\Delta_{bs} \neq 0$) greatly affects inferences about the size, sign and statistical significance of estimates of schooling quality on the return to schooling as we note below.

IV. Empirical Findings

We now report estimates of the effect of quality on earnings obtained from the Johnson-Stafford and Card-Krueger models discussed in the previous section. We test whether restrictions maintained in these models are consistent with the data and find that many are not. In particular, we reject the assumption maintained in both models, and in most of the schooling quality literature, that log earnings is a linear function of education. Significant "sheepskin" effects, especially for graduation from college, give rise to nonlinearities in the return to education. When the earnings-schooling model is modified to allow for these nonlinearities, support for the quality hypothesis weakens.

Another assumption maintained in both models that we test and reject is that there are no interactions between region of birth and region of residence (see JS-2 and CK-1, CK-2). The significant interaction effects that we find, both in levels and in slopes, are consistent with the evidence presented in section III on the importance of selective migration. Individuals from particular regions of birth tend to migrate to particular

---

[13] See, e.g., Heckman and Scheinkman (1987) for discussions of how non-uniform pricing of attributes across sectors in an economy gives rise to the problem of self-selection. They present empirical evidence against uniform pricing in the U.S. economy.

regions of residence, presumably because they have a comparative advantage in certain labor markets possibly arising from lower migration costs or higher expected earnings gains. When interactions between region of residence and regions of birth are introduced into the earnings model to account for patterns of comparative advantage, estimated quality effects change considerably. Estimates obtained from this model also indicate that a unit increase in quality has a different effect on the rate of return to schooling in different labor market regions.

A Model with Quality Entering Through Intercepts

The first specification we examine is that of Johnson and Stafford (1973), in which quality is assumed to have a uniform effect on log earnings, regardless of the educational attainment level. In the notation from the last section, this model is written as:

$$(14) \qquad y_{ibsc} = \theta_c + X_{ibsc} \cdot \beta_c + E_{ibc} \cdot \alpha_c + \delta_c \ln(Q_{bc}) + \varepsilon_{ibsc} \, ,$$

where X_{ibsc} is a matrix with demographic variables for individual i born in state b in cohort c living in state s. E_{ibs} represents years of education, Q_{bc} a vector of the four quality measures described in section II, and ε_{ibsc} an individual specific disturbance term assumed to be i.i.d. across individuals. X_{ibsc} contains an SMSA indicator, potential experience (experience = age - 6 - schooling), and potential experience squared. This model imposes restrictions JS-1 through JS-3 from the previous section.[14]

In section I, we showed that an earnings specification similar to the JS model can be derived within an efficiency units human capital model, where the amount of human capital possessed by an individual is the product of his or her educational quality and years of schooling. The model we derived (given by (4)) included residence indicators in intercepts to account for varying skill prices across labor markets. These are notably absent from the JS model. Differences in regional labor markets or in costs of living are assumed to be fully captured by an urban residence indicator. Similarly, a person's state

---

[14] Note that our analysis does not replicate the Johnson-Stafford (1973) study. We use a different data set and rely on log weekly wages as opposed to log hourly wages as the dependent variable. Johnson and Stafford's quality measure is per pupil expenditure, whereas our measures are pupil-teacher ratios. term length, and relative teacher salary.

of birth plays a role only through the associated measure of quality. Failure to account for other determinants of earnings that are correlated with schooling quality, such as family background, could lead to bias in the estimated quality coefficients.[15]

Table 7 presents the estimated coefficients for the JS model using 1990 Census data. In these results, and all those that follow, we use the pupil-teacher ratio based on attendance rather than on classroom enrollment and report Eicker-White robust standard errors. Figures 1(a) - 1(b) reveal the difference between these two measures of the pupil-teacher ratio.[16] The top panel of Table 7 displays the estimated coefficients for the model in (14). The bottom panel adds to the model interactions between years of education and the log of the quality measures, so that it informs us of how the effect of schooling quality varies with years of education.

For the base model in the top panel, we find significant effects of term length and teacher salary on earnings for all cohorts. However, the coefficient associated with the pupil-teacher ratio is always perversely signed.[17] Results obtained for cohorts from the 1970 and 1980 Census samples are similar and are therefore omitted for the sake of brevity.

When Johnson and Stafford supplement (14) with an interaction term between log quality and education, they find the associated coefficient to be statistically insignificant and that the inclusion of this interaction has little effect on the estimated quality coefficients. In our data, the interaction is numerically and statistically significant and affects the estimated quality coefficients. It is strongest for the 1980 and 1990 Census years but weaker in 1970, so changing importance over time may explain why Johnson and Stafford find no support for an interaction in the earlier data. The coefficients associated with term length and relative teacher salary indicate that the effect of quality increases with higher levels of education. This evidence provides support for the

---

[15] Johnson and Stafford attempt to control for such bias using an indicator for whether the individual grew up in an urban area and for the father's education level, but we do not have that information.

[16] Figure 1(b) demonstrate the convergence in schooling quality across regions over time, a convergence that has led to the speculation that returns to quality should be less precisely estimated for recent cohorts because of a lack of variance. See Heckman, Layne-Farrar, and Todd, 1995, for a discussion of this point.

[17] Recall that increases in the pupil-teacher ratio represent a decrease in educational quality.

assumption of Behrman and Birdsall (1983) and Card and Krueger (1992b) that quality enters through the rate of return to education. However, model intercepts also appear to be an important source of information about quality.[18]

A Model with Quality Entering Through Slopes

We use the Card and Krueger (CK) model as a starting point for examining the effect of quality on returns to education. In their specification, y depends on a person's state of birth (b), state of residence (s), years of education (E), and several demographic variables (X).[19] In the notation of the last section, this model can be written as

$$(15) \quad y_{isbc} = \tilde{\theta}_{bc} + \tilde{\theta}_{sc} + X_{isbc} \cdot \beta_c + E_{isbc} \cdot (\tilde{\alpha}_{bc} + \tilde{\alpha}_{r(s)c}) + \varepsilon_{isbc}$$

- i refers to the individual, b to the state of birth, s to the state of residence, $r(s)$ to the region of residence, and c to the birth cohort

- E_{isbc} is the number of years of schooling completed

- $\tilde{\alpha}_{bc} = \alpha_c + \alpha_b + \alpha_{bc}$ is the "return" to education for an individual born in state b in cohort c inclusive of the cohort mean return

- $\tilde{\theta}_{bc} = \theta_c + \theta_b + \theta_{bc}$ is a cohort-specific fixed effect for state of birth inclusive of the cohort mean

- $\tilde{\theta}_{sc} = \theta_s + \theta_{sc}$ is a cohort-specific fixed effect for state of residence

---

[18] Johnson and Stafford also do not find support for an interaction between years of education and years of work experience. When we include an interaction between education and a quadratic in experience, we find it to be highly significant, indicating that experience-earnings profiles differ by education level. This contradicts evidence by Mincer(1978) for earlier census years. We plan to pursue this further in other work.

[19] Card and Krueger's definition of education is somewhat unusual: E is the years of education above the level attained by the bottom 2% of people in the education distribution. Education is thus defined by $E_{ibc} = Max(e_{ibc} - T_{bc}, 0)$ where e_{ibc} is the actual level of educational attainment and T_{bc} is the bottom second percentile of the education distribution. For the cohorts we examine, the 2% threshold is typically around 8-9 years of education. Persons with this level or below are assumed to receive no returns to their education under the CK framework. After having determined through sensitivity analysis that assuming linearity in education above the 2% threshold yields similar findings to a model assuming linearity in actual years of education, we decided to use the latter more conventional modeling assumption for our empirical analysis. All of our results are therefore based on actual years of schooling attended.

- X_{ibc} is a matrix of individual characteristics: a marital status indicator, potential work experience, potential work experience squared, and an SMSA indicator, with associated coefficients β_c

- $\tilde{\alpha}_{r(s)c} = \alpha_{r(s)} + \alpha_{r(s)c}$ is a region of residence effect, meant to capture regional labor market shocks to the education return (see below)

- ε_{ibsc} is a stochastic error term assumed to be i.i.d. across individuals.

Quality enters into the earnings function through its effect on the return to education. The CK model posits a linear relationship between the rate of return for individuals born in state b in cohort c, $\tilde{\alpha}_{bc}$, and educational quality, Q_{bc}, as measured by the pupil-teacher ratio, teacher salary and term length:

$$(16) \quad \tilde{\alpha}_{bc} = \alpha_b + \alpha_c + \varphi Q_{bc} \quad ,$$

where α_b is a state of birth-specific intercept term and α_c is a cohort effect.

Estimates of φ are obtained from a two-stage procedure regressing estimates of $\tilde{\alpha}_{bc}$ on Q_{bc}. That is, the first stage equation (15) is estimated separately by cohort, and then the rate of return estimates for all cohorts are combined in the second stage to estimate an effect of quality. If we substitute equation (16) into equation (15), we see that quality only enters the log earnings equation multiplicatively with education. This presumes that the effect of a change in quality on earnings is monotonically increasing with years of education. Also, since quality affects the rate of return to education independently of the region of residence, an implication of the CK model is that within a given education level, rankings of wages by state of birth should be stable across regions of residence. The nonparametric rank tests reported in Section II showed some support for stability in the rankings in the 1990 data, and less support in other Census years.

Several other features of this specification are noteworthy. First, observe that the parameterization of work experience (included in X_{ibsc}) is not symmetric with respect to the parameterization of the effect of education. Since work experience is a form of human capital that is complementary with formal schooling and since participation in formal school increases the rate of accumulation of job training (Mincer, 1974), it is not clear that

an asymmetric treatment of experience is warranted. We test this assumption and find evidence that experience is valued differently in different regions (i.e., $\beta_{sbc} \neq \beta_c$). When experience is introduced into the earnings equation symmetrically with education, there is little evidence that schooling quality has an effect through the return to work experience. However, relaxing the restriction on the returns to experience leads to much weaker estimated effects of quality operating through the returns to education.

Second, this specification assumes that the earnings-schooling relationship is linear. Below we test -- and reject -- this specification for all three Census years. Estimates from a more flexible model that allows for nonlinearities indicate that the only support for an effect of quality on the rate of return to schooling comes through the marginal returns to attending or completing college. For other education levels, there is no evidence for an effect of quality.

Third, equation (15) rules out state of birth - region of residence interactions in both slopes and intercepts ($\theta_{br(s)}$=0 and $\alpha_{br(s)}$=0). As shown in section III, migrants' wages are typically higher than those of nonmigrants, which indicates that migration is probably selective on wages in the destination state. Since the model in (15) assumes that ε_{ibsc} is exogenous with respect to the current location indictors D_{si} (the variables associated with $\theta_{r(s)c}$ and $\alpha_{r(s)c}$), it does not allow for the possibility of selective migration. When interactions between regions of birth, $r(b)$, and regions of residence, $r(s)$, are included to account for possible endogeneity, estimated schooling quality effects often reverse signs or become imprecisely estimated.

Fourth, this specification assumes that returns to education are region-specific ($\alpha_{r(s)c} \neq 0$). We test and do not reject this hypothesis in the context of a linear model. The introduction of the regional deviation term, $\alpha_{r(s)c}$, is an important contribution of the Card and Krueger analysis. This term is intended to allow for regional and cohort-specific labor market shocks in the return to education. Marginal returns to education, used in the second stage, are therefore estimated relative to region-specific means. The justification for estimating "returns to education" ($\tilde{\alpha}_{bc}$) relative to a regional mean depends on the source of the shock captured in $\alpha_{r(s)c}$. If the high value of $\alpha_{r(s)c}$ is determined in part by

regional school quality aggregates ("supply-side shocks"), then the total contribution of quality to the marginal returns to education is overstated, because higher quality reduces marginal returns. In fact, quality might impact earnings through a variety of different channels, including its effect on marginal returns to education and its effects on regional prices operating through aggregates. We find evidence that regional quality aggregates should be included in earnings equations. The specification of equations (15) and (16) implicitly assumes that cohort and region-specific shocks are demand-driven or are unrelated to quality variation within regions, but our analysis reveals an important role for aggregate supply shocks.

In the empirical results that follow, we extend the CK framework by permitting $\alpha_{r(s)}$ and $\alpha_{r(s)c}$ (the regional components of the return to education) to depend on regional demand and supply aggregates, such as the aggregate level of schooling quality in the local labor market and the percent of the labor market's workforce employed in manufacturing, in order to produce a richer economic model of the labor market pricing equation in the tradition of Tinbergen (1975) and Sattinger (1980). Denoting the vector of aggregates by Z, our extension parameterizes the region of residence components in the following way:

$$\alpha_{rc} = Z_{rc} \pi$$

$$\alpha_r = Z_{r.} \pi$$

It is through these channels that the first-order effects of economic aggregates should flow. We find that quality aggregates usually operate like supply shocks and depress the rate of return to schooling. Measuring the effect of quality on earnings solely through its effect on α_{bc} frequently overstates the contribution of quality to earnings.

Estimated Impacts of Schooling Quality

From equation (15), we obtain estimates of the total return to education $(\tilde{\alpha}_{r(s)bc} = \tilde{\alpha}_{bc} + \tilde{\alpha}_{r(s)c})$ for each (b,s,c) cell. These returns are then combined across cohorts and regressed on the quality variables in the second stage,[20]

---

[20] This procedure differs from that used by Card and Krueger in that the total return is used as the dependent variable in the second stage instead of only the state of birth component of the return. As shown in the last section, Card and Krueger's procedure fails to deviate the independent variables in the

$$\tilde{\alpha}_{r(s)bc} = \alpha_c + \alpha_b + Q_{bc}\varphi + \alpha_{r(s)} + \alpha_{r(s)c} \quad ,$$

where α_c , α_b , α_{rc} , α_{rc} are fixed effects for cohort, state of birth, state of residence, and state of residence interacted with cohort.

Table 8a, columns 1-3, reports estimates of the effect of quality, Q, on the state of birth components of the rate of return ($\tilde{\alpha}_{bc}$) using 1980 Census data. The only difference between column 2 and column 1 is the addition of two additional schooling quality measures. Column 3 is like column 1 but the state of birth fixed effect, α_b , is suppressed so that there is a single intercept term. These first three columns present the specifications closest to those reported by Card and Krueger. They provide support for the schooling quality hypothesis: the estimated quality coefficients are statistically significant and have signs consistent with the hypothesis that quality raises the return to education.

Despite the differences between our model and Card and Krueger's original specification, we replicate their basic findings using data from two additional Census years.[21] Results for 1970 and 1990 are reported in Tables 8b and 8c respectively. (See column 1 in each table). The pattern of strong quality effects found in 1980 holds up in 1970 and in 1990. The 1990 results are even stronger (in absolute value terms) and statistically more significant. Column 2 adds two other measures of schooling quality which, are also found to be statistically significant in all Census years. The pattern of amplification of estimated quality effects in 1990 is consistent with other evidence that the returns to schooling and ability have increased between 1980 and 1990. Column 3 shows generally weaker estimated quality effects in all years. In 1980 and 1990 the estimated

---

second stage from regional means and is therefore only valid when regional means and national means coincide. By using the total return to education as the dependent variable and including regional indicators in the regression , we deviate both the dependent and independent variables from regional means.

[21] A comparison of the estimates in Table 7a with Card and Krueger's (1992b) estimates (reported in their Table 1) shows that deviating around regional means in the second stage has little impact on the coefficients Also, using days attended rather than days enrolled weakens - but not eliminate - the estimated impact of the pupil teacher ratio on earnings.Using their quality data and their precise specification, we were able to replicate Card and Krueger's empirical findings almost exactly.

39

pupil-teacher ratio coefficients become statistically insignificant. The results do not appear robust to the inclusion or exclusion of α_b from the model.[22]

To interpret the estimated quality coefficients, we present Table 9, showing the effects of hypothetical improvements in schooling quality on earnings. No single quality variable emerges as the most important for effecting a change in earnings. The predicted increases in earnings resulting from changes in the teacher wage and term length measures are relatively larger in 1990.

Table 9
Effect of Hypothetical Improvements in Quality
on Earnings for High School Graduates and College Graduates
Model: Extension of Linear Model (coefficients given in tables 8a-c, column 2)

Change in Quality	1970		1980		1990	
	12 years	16 years	12 years	16 years	12 years	16 years
Decrease pupil-teacher ratio (attendance based measure) by 5	↑5.33%	↑7.12%	↑3.76%	↑5.01%	↑6.40%	↑8.53%
Increase relative teacher wage by 30%	↑4.32%	↑5.76%	↑7.49%	↑9.98%	↑9.32%	↑12.4%
Increase term length by 10 days	↑0.89%	↑1.18%	↑1.28%	↑1.71%	↑7.74%	↑10.3%

(1) Because returns are assumed to be linear in quality, the effects of other magnitudes of
    quality changes are easily found by rescaling the numbers in the table.

Column 4 of Tables 8a-c allows the effect of schooling quality to differ across cohorts ($\varphi = \varphi_c$). For all the Census years, the estimated effect of the pupil-teacher ratio weakens and sometimes becomes perverse for the recent cohorts. This evidence supports the argument of diminishing marginal returns to improvements in quality and declining variability in schooling quality in recent cohorts producing lower estimated returns that are less precisely estimated. However, variability in term length has also declined greatly over time and its estimated effect is strongest for more recent cohorts.

---

[22] This evidence is at odds with findings reported by Card and Krueger (1992b).

40

Columns 5 through 7 in Tables 8a-c extend previous analyses by allowing for additional channels by which quality can affect earnings. State of birth and region of residence components of the total return to education (α_{bc} and α_r, α_{rc}) are parameterized as functions of state of birth quality and regional aggregate quality. Fixed effects for state of birth are included in the regression. The second stage regression is thus given by

$$\tilde{\alpha}_{r(s)r(b)c} = \alpha_c + \alpha_b + Z_{r(s)c}\pi + Q_{bc}\varphi .$$

Columns 5 and 6 demonstrate that regional aggregate quality operates through first-order region of residence market effects ($\alpha_{r(s)} + \alpha_{r(s)c}$). In all three years, the data are consistent with a supply-shift hypothesis. In column 5, higher stocks of quality, as measured by the regional average pupil-teacher ratio, reduce the marginal returns to education. Column 6 includes two aggregate quality measures: pupil-teacher ratio and term length.[23] The coefficents show that higher aggregate quality produces lower rates of return, although the estimate for the aggregate pupil-teacher ratio weakens when the aggregate term length is included.

Column 7 investigates how alternative parameterizations of regional variables affect the estimates of quality operating through a_{bc}, $a_{r(s)}$ and $a_{r(s)c}$. It extends the list of regional aggregates to include the distribution of persons among educational categories and the percent of the workforce in manufacturing interpreted as a demand variable. Now the coefficients on the quality variables are only partly consistent with the story that the aggregates act like a supply shifters (for example, regional average pupil-teacher ratio has the "wrong" sign.) In 1970 and 1980, increases in the proportion of the labor force in the middle range of schooling (high school and 1-3 years of college) depress returns at the mean, while increases in dropouts, interpreted as a reduction in the aggregate stock of skills, raise the return. This is consistent with complementarity across different skill groups. An increase in aggregates of groups near the mean of the skill distribution depresses mean returns. Except for 1980, increases in the percentage of the workforce in manufacturing reduce the returns to schooling. This is consistent with the notion that

---

[23] The regional average of relative teacher salary is equal to one by construction since our measure of teacher salaries is relative to the regional mean. This variable is therefore omitted form the regional aggregate schooling quality measures.

demand for low-skill high wage manufacturing jobs raises the demand for low skill labor. Specifications 5 through 7 also show that estimated quality effects operating through α_{bc} are robust to the inclusion of regional aggregates in the model.

The Effect of Quality Operating Through the Intercepts of the Card-Krueger Model

Our analysis of the Johnson and Stafford model indicated that both intercept and slope parameters are potential sources of information about schooling quality's effect on earnings. We therefore investigate the effect of quality operating through the intercepts of the CK model. Card and Krueger explicitly ignore this channel of influence in an attempt to control for "fixed effects" bias arising from omitted state of birth and state of residence specific intercepts. Nevertheless, it is of interest to know whether focusing only on the rate of return to education leads to an over or understatement of the total effect of quality on earnings, since a opposite effect could operate through intercepts.

The equation relating model intercepts to the state of birth quality measures and regional average quality is:

$$(17) \quad \tilde{\theta}_{sbc} = \theta_s + \theta_b + \theta_c + \varphi Q_{bc} + Z_{r(s)c}\pi \ ,$$

In our treatment of the intercept terms, we use state of residence instead of region of residence to index the current location variables. For now, we estimate the effect of the empirical measures of schooling quality on $\theta_b, \theta_{bc}, \theta_s$, and θ_{sc}, under the assumption that $\theta_{sbc} = \theta_{sb} = 0$ (CK-1). Later, the assumption of no interactions between the state of residence and state of birth is relaxed.

Table 10a presents estimated coefficients for the 1980 Census data. Comparable tables for 1970 and 1990 are found in Tables 10b-10c. The specifications given by each of the columns are analogous to those discussed in relation to the slope parameters. Most of the estimated coefficients on the quality variables are statistically significant. In Table 10b, for 1970, the pupil-teacher ratio operating through the intercepts supports the quality hypothesis and reinforces the slope effect seen earlier in Table 8c. In 1980 and 1990, however, the effect of the pupil-teacher ratio operating through the intercept sometimes reinforces and sometimes counters the effect, depending on the model specification. For the term length and teacher salary variables, we find evidence in all three census years that

the effect of quality on earnings operates through the intercept parameters in a direction opposite to its effect through the slope coefficients. Hence, focusing only on an estimated quality effect operating through slope coefficients leads to an <u>overstatement</u> of the effect of the quality variables on earnings. Specification four, which allows the quality coefficients to differ across cohorts, generally supports these conclusions.

In columns five through seven, the state of residence intercepts are parameterized to be functions of regional aggregates. The signs of the regional average quality measures $(Q_{r(s)})$ are affected by whether other regional aggregates are included in the model. They are not always consistent across Census years. The coefficients of the state of birth quality variables (Q_{bc}) are robust to the inclusion of these aggregates.

The Parameterization of the Work Experience Variable

In the CK model, the coefficients on experience and its square, β_c, vary only by cohort, while the coefficient on education is allowed to vary over states of birth, b, and regions of residence, r. In research not reported here, we treat experience symmetrically and examine the effect of quality operating through the returns to experience. The second stage regresses the total return to experience, calculated at the mean level of experience for each cohort, on the state of birth quality variables.[24] The resulting coefficients associated with the quality measures are generally insignificant, providing little evidence of an effect of quality on earnings operating through the returns to experience. Moreover, when experience is introduced symmetrically, we find that there are no longer significant effects of quality operating through the returns to education. These results are available on request from the authors.

To summarize, we have investigated five distinct channels through which quality can affect earnings: state of birth intercepts, state of residence intercepts, state of birth specific returns to education, regional returns to education, and returns to experience.

---

[24] Specifically, if β_{b1} is the coefficient on experience interacted with state of birth and β_{b2} is the coefficient on its square, we calculate $B_{bc} = \beta_{b1} + \beta_{b2}*$(average experience). B_{bc} becomes the dependent variable in the second stage regressions.

Schooling quality appears to affect the first four of these channels but not always in the same way. There is evidence that quality effects operating through slope parameters frequently overstate the impact of schooling quality impact on earnings, because a reverse effect operates through model intercepts or through the regional return to education. Therefore, pricing equations that limit the scope for quality effects likely misstate the contribution of schooling quality to earnings. Finally, we find no evidence that schooling quality affects the return to post-school work experience. A symmetric treatment of the education and experience variables leads to a weakening of the estimated quality effects on the rate of return to education.

"Sheepskin Effects" in the Return to Education

A critical assumption which is standard throughout the literature is that log earnings equations are linear in years of education.[25] The linear model does not allow for "sheepskin effects," which might be expected to occur upon completion of a high school or college degree. Furthermore, linearity can impose a severe limitation in the way quality is allowed to affect earnings. In the model given by (15) and (16), quality enters only interactively with education, imposing the restriction that a change in quality has a progressively larger effect at higher levels of education.

We test the linearity assumption by adding "sheepskin" effects (dummy variables for grade completion) at grades 8, 12, and 16 to equation (15). F-tests on the significance of the sheepskin effects are reported in Table 11.[26] The data strongly reject linearity for most cohorts in the three Census samples and show that sheepskin effects are important in characterizing the return to education. To illustrate the magnitude of the sheepskin coefficients, they are plotted for the 1980 Census, 1930-39 cohort in Figure 2. Because the Census sample sizes are large, there is a tendency to reject the null hypothesis of no sheepskin effects. Therefore, we show the significance of the coefficients at the 0.1%

---

[25] An exception is a study by Wachtel (1976), in which the quality relationship is estimated separately by educational attainment level.

[26] The F-tests are joint across all states of birth for each cohort at each level of schooling (8th and 12th grades and 4 years of college).

44

significance level ("+" indicates significant at this level). There is clear evidence of discrete jumps in the return to education, especially for completion of a college degree. Thus there is no support for the linearity hypothesis, a finding consistent with previous results in Solon and Hungerford (1987).

A Nonlinear Model

We next revise the earnings model in (15) to allow for nonlinearities in the log earnings equation. Given the evidence of especially significant sheepskin effects at the college level, we estimate a model that allows for discrete jumps in the return to education at 1-3 years of college and at 4 or more years of college.[27] Since the samples we use are very large, it may be the case that the statistical rejections reported above are of little economic consequence and that linearity may not be grossly at odds with the data. If this were the case, relaxing the linearity assumption should have only a slight impact on estimates of the effect of quality on schooling. Instead, we find that it has a significant impact.

The nonlinear model can be written as:

$$(18) \qquad y_{ibsc} = \tilde{\theta}_{sbc} + \tilde{\alpha}^0_{r(s)bc} \cdot e_{isbc} + \tilde{\alpha}^1_{r(s)bc} \cdot \tau^1_{isbc} + \tilde{\alpha}^2_{r(s)bc} \cdot \tau^2_{isbc} + X_{ibsc} \cdot \beta_c + \varepsilon_{ibsc}$$

where

$$\tilde{\alpha}^j_{r(s)bc} = \alpha^j_c + \alpha^j_b + \alpha^j_{r(s)} + \alpha^j_{bc} + \alpha^j_{r(s)c} \quad , \quad j = 0,1,2$$

$$\tilde{\theta}_{sbc} = \theta_c + \theta_b + \theta_s + \theta_{sc} + \theta_{bc}$$

$$\tau^1_{isbc} = \begin{cases} 1 & \text{if individual i completed some college but less than 4 years} \\ 0 & \text{else} \end{cases}$$

$$\tau^2_{isbc} = \begin{cases} 1 & \text{if individual i completed 4 or more years of college} \\ 0 & \text{else} \end{cases} \quad ,$$

---

[27] Since each additional education category adds 76-120 parameters to the model (depending on the specification), we did not introduce more than two marginal returns. Even though the nonlinear model is highly parameterized, our sample sizes are large (see Appendix A) and the coefficients are still precisely determined.

and where e_{isbc} is actual years of education. X_{isbc} represents the same demographic variables as used in the linear model (defined under equation (15)).

Since $\alpha^0_{r(s)bc}$ represents the per-year return to school, it is multiplied by the years of education to obtain the total return to schooling for someone with a high school degree or less. For someone with 13 years of school, the total return to education would be the linear effect plus the marginal effect, $\alpha^0_{r(s)bc}*13 + \alpha^1_{r(s)bc}$. If an individual has a Master's degree (assumed to be 18 years of school), his or her return is given by $\alpha^0_{r(s)bc}*18 + \alpha^2_{r(s)bc}$. To illustrate the magnitude of the estimated education returns for the model given in (8), Figure 3 plots them for a few representative states.

From equation (18), we obtain estimates of the linear base rate of return ($\tilde{\alpha}^0_{r(s)bc}$) and of the two marginal returns for some college and for four or more years of college ($\tilde{\alpha}^1_{r(s)bc}$ and $\tilde{\alpha}^2_{r(s)bc}$). These estimates are then used in a second stage regression, in which we parameterize the different components of the return as functions of quality. The second stage equation is estimated separately for the linear base return and for each of the marginal returns. As with the linear model, we consider the effect of quality in the state of birth operating through the state of birth components, α^j_b and α^j_{bc}, as well as the effect of aggregate quality on regional labor market conditions, operating through region of residence components, $\alpha^j_{r(s)}$ and $\alpha^j_{r(s)c}$.

Tables 12(a-1) - 12(a-3) present estimates of the effect of quality on the rate of return where the dependent variables are the linear base return, the marginal return for some college, and the marginal return for four or more years of college. The procedure we follow is analogous to that described for the linear case and the estimated coefficients for the nonlinear model can be compared to those for the linear model in Table 8(a). In specifications three and four of Table 12(a-1), which parameterize α^j_b and α^j_{bc} as functions of quality, the signs of the quality coefficients are often reversed or else become statistically insignificant. Interestingly, support for an effect of secondary school quality on earnings comes through the marginal return for some college or for a college degree (see Tables 12(a-2), 12(a-3)). This finding is consistent with results reported by Wachtel (1976). As previously seen with the linear model, estimated quality effects differ in

magnitude depending on whether fixed effects for state of birth are included in the model. The final three specifications reported in the tables examine the effect of quality operating through the regional return, $\alpha^j_{r(s)}$ and $\alpha^j_{r(s)c}$. Some statistically significant effects appear, but they are not interpretable because of many sign reversals.

The second stage regression results for 1970 and 1990 Census data are given in Appendix C in tables 12(b-1)-12(b-3) and 12(c-1)-12(c-3). For 1970, there is some evidence of an effect of term length on the base return in specification (2), but the signs of the other estimated quality coefficients are generally inconsistent with the quality hypothesis. Support for an effect of quality on earnings again comes mainly through the marginal return to attending or completing college. In 1990, the pupil-teacher ratio is consistently of the expected sign only when the dependent variable is the marginal return to completing 4 or more years of college.

Relaxing the false assumption of linearity produces weak support for an effect of quality on earnings for individuals who attend college. Since students from higher quality schools often attend better colleges, in which case part of the estimated effect of secondary schooling quality is more properly attributable to college quality, a variable omitted from our regression.

To illustrate the effects of schooling quality on earnings implied by the estimated coefficients, we tabulate the effects of counterfactual changes in quality in Table 13. The results are unstable across measures and years and do not provide strong support for the quality hypothesis.

Table 13
Effect of Hypothetical Improvements in Quality on Earnings
Nonlinear Model
(Coefficients in Tables 12(all), column 2)

Change in Quality	1970		1980		1990	
	12 years	16 years	12 years	16 years	12 years	16 years
Decrease pupil-teacher ratio (attendance based measure) by 5	↓1.39%	↑0.89%	↑1.77%	↑3.77%	↓5.08%	↓2.21%
Increase relative teacher wage by 30%	↑0.79%	↑1.81%	↑3.71%	↑7.17%	↑6.16%	↑6.95%
Increase term length by 10 days	↑3.24%	↑2.00%	↓2.27%	↓1.23%	↑5.06%	↑7.14%

*Accounting For Comparative Advantage and Region of Birth * Region of Residence Interactions*

The empirical results presented thus far are predicated on the absence of any interactions between region of residence and state of birth (θ_{rb}=0 and $\alpha_{r(s)r(b)}$=0, given by assumptions JS-1, CK-1 and CK-2). Yet, the evidence in section II indicates that individuals from particular states of birth tend to migrate to particular states of residence, presumably because they have a comparative advantage in certain labor markets. If their expected earnings gains from migration are correlated with realized values of the disturbance term in the earnings equation, ε_{ibsc} , then the disturbances are no longer exogenous with respect to the regressors.

To allow for patterns of comparative advantage in particular geographic regions, we relax the assumptions that $\tilde{\theta}_{r(s)r(b)c} = 0$ and $\tilde{\alpha}_{r(s)r(b)c} = 0$. For the sake of brevity, the tables display only the empirical estimates for the linear model, because the nonlinear model leads to similar conclusions regarding the instability of estimated effects of quality when the model includes interaction variables.[28] Also, in the interest of fewer model parameters, interactions are introduced into the linear model in terms of regions of birth

---

[28] The results for the nonlinear version of the model are available from the authors.

and regions of residence (r(b) and r(s)) instead of states. F-tests of the assumption of no interactions, reported in Table 14, decisively reject this restriction.

We introduce interactions between region of residence and region of birth into the model stated in (15) both in intercepts and in slopes:

$$(19) \qquad y_{isbc} = \tilde{\theta}_{sbc} + X_{isbc} \cdot \tilde{\beta}_c + E_{isbc} \cdot \tilde{\alpha}_{r(s)bc} + \varepsilon_{isbc}$$

$$\tilde{\theta}_{sbc} = \theta_s + \theta_b + \theta_c + \theta_{bc} + \theta_{sc} + \theta_{r(s)r(b)} + \theta_{r(s)r(b)c}$$

$$\tilde{\alpha}_{r(s)bc} = \alpha_{r(s)} + \alpha_b + \alpha_c + \alpha_{bc} + \alpha_{r(s)c} + \alpha_{r(s)r(b)} + \alpha_{r(s)r(b)c}$$

where $I = 1,...,I,\ s = 1,...,S\ ;\ b = 1,...,B,\quad c = 1,...,C\quad,r = 1,...,R$.

$\tilde{\beta}_c$ is constrained to be the same across states but is permitted to vary across cohorts. First, estimates of the total return to education, $\tilde{\alpha}_{r(s)bc}$, are obtained for each (b,s,c) cell. The estimates for all the cohorts are then combined and, in a second stage regression, components of the return to education are parameterized as functions of quality.

In the most general treatment of the interaction terms $(\alpha_{r(s)b}\ ,\alpha_{r(s)bc})$, the quality effect operating through α_b and α_{bc} could be estimated without imposing restrictions on the interactions. However, a problem that arises in this specification, and in any model with interactions, is that the estimated main effects will generally depend on which interactions are omitted as the base categories in the regression.

To illustrate this point, consider a simple model in which the outcome Y can be described by two main effects and one interaction. Assume that there are two possible states associated with each main effect: $d_1=(0 \text{ or } 1)$ for the first and $d_2=(0 \text{ or } 1)$ for the second. One way to write the outcome equation selects $d_1=0$ and $d_2=0$ as the base categories:

$$Y = \mu + \alpha\, d_1 + \beta\, d_2 + \gamma\, d_1 d_2 \ .$$

Of course, selecting $d_1=1$ and $d_2=0$ as the base categories would be an equally valid representation which would yield different definitions of the main effects,

$$Y = \mu^* + \alpha^*\, d_1 + \beta^* d_2 + \gamma^*(1 - d_1)d_2 \ ,$$

where $\alpha^*=\alpha$, $\beta^* = \beta - \gamma$, and $\gamma^* = -\gamma$.

Because of potential sensitivity to the chosen basis, we estimate the second stage model, in which the dependent variable is the total return to education, using a variety of base categories. A general result obtained from these unrestricted interaction models is that there is much less support for an effect of quality on earnings once the assumption of no interactions is relaxed. Table 15a presents estimates obtained for the linear model using 1980 Census data, where the benchmark region of residence and region of birth is East North Central. Comparable estimates for 1970 and 1990 data are found in Appendix C, Tables 15b-15c.

In the first three columns, the state of birth components of the return to education are parameterized as functions of quality. A comparison with estimates in Table 8a for the linear model without the interaction terms reveals that the effect of including the interaction terms is to greatly diminish the importance of our schooling quality proxies in the earnings equation. The coefficient on the pupil-teacher ratio is almost always perversely signed or insignificant. The estimated effect of term length is only significant for the third specification in which the fixed effects for state of birth are suppressed, while teacher salary is significant only when state of birth fixed effects are included. When the quality effects are allowed to vary across cohorts, most of the effects are insignificant. The estimated coefficients in the 1970 data lead to identical conclusions. In 1990, the pupil-teacher ratio is significant only in those specifications that include state of birth fixed effects, and the term length variable only in the specification that excludes fixed effects. There is some support for an effect of relative teacher salary that is robust to whether or not the model includes state of birth intercepts.

To look further into the determinants of the significant interaction patterns, we parameterize the interaction terms in simple and interpretable ways. We relax the assumption maintained in all the models considered thus far that the effect of a unit increase in schooling quality on the rate of return to education is the same across regions. We also allow the rate of return to education for individuals born in state b living in region r to depend on the physical distance between regions. Since longer distances traveled are usually associated with higher costs of migration (both actual and psychological costs), we

would expect the rate of return to increase with the distance of the move. Formally, we write,

$$(20) \quad \alpha_{r(s)bc} = \alpha_b + \alpha_c + \alpha_{r(s)} + \alpha_{r(s)c} + \varphi_{r(s)} Q_{bc} + \eta_1 d_{r(s)r(b)} + \eta_2 d^2_{r(b)r(s)}$$

where φ_r varies across regions and $d_{r(s)r(b)}$ is the distance between regions of birth and regions of residence. As noted above, Card and Krueger make the assumption that $\varphi_r = \varphi$ for all regions and ignore any effects of distance on returns to schooling. Although the model in (20) relaxes assumptions made in the literature, we find evidence that it may still be too simple. When we compare the model with unrestricted interactions to the model in which the interactions are parameterized using the Schwartz model selection criterion, the unrestricted model is favored (see Table 16).

Estimated coefficients on the quality variables in model (20) indicate that unit increases in quality generally have different effects on the rate of return to schooling in different regions. The estimates shown in Table 17a are for a model with linear schooling. The specifications in the first three columns are for three different quality measures, taken one at a time, in a model that includes fixed effects for the state of birth. In most regions, a lower pupil-teacher ratio is associated with a higher rate of return to schooling, but this is not the case in New England, where a higher pupil-teacher ratio raises the returns to schooling. The coefficient on distance and its square is highly significant and in the expected direction in all specifications. Higher distance implies higher wages. As seen by columns (2) and (3), relative teacher salary has a similar effect on the rate of return across regions, but estimated effects of term length are always insignificant or perversely signed.

Columns (4) and (5) consider all three quality measures in specifications with and without fixed effects for state of birth. As was found in the previous models, estimated coefficients are not robust to the exclusion of these fixed effects. In column (5), the coefficient on pupil-teacher ratio is consistently of the wrong sign and the coefficient on teacher salary is diminished in magnitude and significance. The significant coefficients on term length are generally positive, but again the effect in New England runs counter to that in other regions and counter to the quality hypothesis.

Tables 17b and 17c contain results for the 1970 and 1990 Census years. In models with state of birth intercepts, the pupil-teacher ratio and teacher salary measures show

significant effects while the effect of term length is usually insignificant or in the wrong direction. The strongest effects for the first two quality variables are seen in the 1990 data. Excluding state of birth effects (column (5)) generally leads to a decrease in the estimated pupil-teacher ratio and teacher salary effects, with some sign reversals for the pupil-teacher ratio in 1970. The exclusion leads to an increase in the estimated effect of term length, which becomes strongly significant for most regions in 1990.

Table 18 tests for each Census year the restriction that the coefficients associated with the quality variables are identical across regions ($\varphi = \varphi_r$). This assumption, maintained in all the models considered earlier and maintained in the previous literature, is rejected for the pupil-teacher ratio and term length but is not rejected for teacher salary. Therefore, unit increases in quality generally have different effects on the rate of return to schooling in different regions. This could arise because of differential regional demand factors, or for other reasons previously discussed regarding comparative advantage and selective migration. Thus our evidence speaks against the crucial identifying assumptions used in previous studies. When the assumptions are relaxed, the empirical findings challenge a story that schooling quality raises earnings.

The Effect of Schooling Quality on Educational Attainment Levels

Thus far, we have only considered the effect of schooling quality on earnings, holding constant the level of educational attainment. Another way for quality to raise earnings is by increasing the incentive for students to further their education. Students attending better secondary schools may be more likely to graduate from high school or more likely to attend college. This result has been found in several previous studies.[29] (Appendix A presents a model that demonstrates how higher quality of schooling produces higher levels of schooling attainment.)

To address this question, we estimate a log linear relationship between the relative number of college and high school graduates in each state to the quality indicators:

(20)
$$\ln[N_{jbc}/N_{hbc}] = \alpha_c + \ln[Q_{bc}] \cdot \beta_c$$

---

[29] Both Card and Krueger (1992b) and Johnson and Stafford (1973) find evidence for this effect.

c designates the cohort and b the state of birth. N_{jbc} is the total number of individuals born in state b in cohort c who report "j" as their highest level of education. N_{kbc} is defined analogously. In the first set of regressions, N_{jbc}/N_{kbc} is the number of college graduates relative to all high school graduates with less than four years of college, ($e <$ 16), for each state b and cohort c. In the second set, j refers to high school graduates with less than four years of college (again $e < 16$) and k to high school dropouts (i.e., $e < 12$). Table 19 tabulates the national means of these ratios for each cohort and each Census year.

<div align="center">

Table 19
Relative Numbers of College and High School Graduates
US Averages Over Time

</div>

Ratio N_{jc}/N_{kc}	Census Year:	Cohort				
		1910-19	**1920-29**	**1930-39**	**1940-49**	**1950-59**
College to	1970	0.28	0.37	0.40	0.26	...
High	1980	...	0.42	0.45	0.56	...
School	1990	...	...	0.44	0.54	0.44
High	1970	0.82	1.16	1.89	3.51	...
School to	1980	...	1.48	2.59	4.61	...
Drop Out	1990	...	...	3.28	5.94	9.09

The ratios show considerable variation across states. East South Central, which consists of Alabama, Kentucky, and Tennessee, has the lowest percentage of college graduates over the entire time period. In 1970 data, only 8% of the 1910-19 cohort born in that region received a college degree or higher and fully 59% were high school dropouts. In 1990 data, 22% of the 1950-59 cohort graduated from college and 14% dropped out of high school. The Pacific region, with California, Oregon, and Washington, and the Mid-Atlantic region, with New York, New Jersey, and Pennsylvania produce the highest percentages of college graduates. Around 15% of the 1910-19 cohort born in these regions have college degrees according to 1970 Census data. This number increases to 28% for the Pacific region and 35% for the Mid-Atlantic region for the 1950-59 cohort in 1990 Census data.

We estimate equation (20) by regressing the log of the total number of college graduates relative to high school graduates on the log of the four quality measures[30] Table 20 displays the coefficient estimates when the dependent variable is the high school graduate to dropout ratio. It is consistently positively related to most of the schooling quality indices for all cohorts. For example, if average classroom size decreased by 1% for the 1910-19 cohort (in 1970 data), the coefficients imply a 1.29% increase in the average number of high school graduates relative to dropouts. The ratio of college to high school graduates is similarly positively related to some of the quality measures, although the relationship is weaker.[31]

This evidence provides some support for the hypothesis of a positive effect of schooling quality on schooling attainment previously found in the literature. The relationship is robust across Census years but may be spurious for reasons that Card and Krueger enunciate in their discussion of earlier studies of the earnings-quality relationship: estimated quality effects are confounded with family background and state of residence effects on the schooling decision. Variations across states in educational attainment levels could be due, in part, to differences in family backgrounds or to different local labor market conditions. The strategy of using worker mobility to identify quality effects on earnings cannot be applied to estimate the quality-educational attainment relationship, because there are no comparable mobility data on students across school systems.

V.   The Endogenous Determination of Schooling Quality

The theoretical models in section I and the empirical models in sections III and IV all assume that schooling quality is exogenously determined. Given a well-established link between family background and residential location decisions, this assumption is clearly questionable. In this section, we investigate the consequences of relaxing the assumption on the estimated quality effects. We develop a model of collective decision making that

---

[30] We also estimated a version of the model where the dependent variable is the log of the proportion of individuals in each state reaching educational level j $\ln(N_{jsc}/Population_{sc})$. The results and conclusions were similar. See Heckman, Layne-Farrar, Todd (1995).

[31] Complete results are available from the authors.

can explain the peculiar finding in the literature that studies based on aggregated quality data tend to find stronger effects than studies using school or district level data. This observation is at odds with the usual criticism directed against using individual level data that estimates of the effect of quality are subject to sorting bias. If econometric techniques are inadequate for controlling the bias, then micro estimates should be stronger, not weaker, than the estimates of quality effects based on aggregate schooling quality measures.

To explain why estimated aggregate effects are stronger, several authors argue that there may be measurement error in the inputs that is averaged out in state aggregates (Card and Krueger (1994), Betts (1995), Hanushek (1994)). While this explanation is logically possible, averaging the quality data could also substantially increase sampling variability by eliminating true variation in the data. The stronger estimated effects estimated on aggregate data are also not consistent with a story of choices by schools and parents within a standard common coefficient linear in parameters educational production function, unless redistribution across districts is *against* high-achieving students. If high-achieving students receive more resources, and resources raise outcomes, averaging should *weaken* and not strengthen estimated input-output relationships.

In section IV, we demonstrated that the functional form assumed for the earnings function plays a strong role in establishing quality effects using aggregated data. When functional form restrictions that are rejected by the data are relaxed, estimated quality effects weaken. In this section, we offer an alternative way of reconciling weak effects reported when disaggregated quality data are used in micro earnings functions and strong effects reported when aggregated data are used in micro earnings functions. Our model shows how maximizing behavior by social planners in allocating schooling funding engineers a dependence between the return to education and the aggregation level of the schooling quality variable, leading to a bias in aggregate data studies. The direction of the bias in estimated quality effects depends on whether funds are allocated in a utilitarian or egalitarian way.

Let the outcome of person i in year t in school s in district d, Y_{itsd}, be written as

$$(21) \qquad Y_{itsd} = \alpha_{sd} + \beta_{sd} Q_{tsd} + \varepsilon_{itsd}$$

where Q_{tsd} is quality at time t in school s in district d. Quality is assumed to be allocated independently of individual performance (ε_{istd}).

Suppose that social planners - or school districts - allocate funds on the utilitarian principle of maximizing aggregate return. Those schools within a district with higher values of β_{sd} would have higher levels of Q_{tsd} i.e. there should be positive regression dependence between β_{sd} and Q_{tsd}.[32] (Egalitarian equalizers would engineer a negative dependence between Q_{tsd} and β_{sd}.)

Using district-level data to proxy schooling quality would produce an upward bias for the average value of β_{sd} within district d. Aggregating up geographical units could substantially *overstate* the average effect of quality on outcomes.

More precisely, assume that the β_{sd} are realizations from a distribution with conditional mean $\beta_{.d}$:

$$E(\beta_{sd} \mid d) = \beta_{.d} \quad ,$$

where "E" denotes mathematical expectation. We may write

$$E(Q_{tsd} \mid d) = Q_{t.d}$$

and $$E(\alpha_{sd} \mid d) = \alpha_{.d}.$$

Using these results, we may write

$$
\begin{aligned}
Y_{itsd} = \alpha_{.d} &+ \beta_{.d} Q_{t.d} + \{\varepsilon_{itsd} + (\alpha_{sd} - \alpha_{.d}) \\
&+ (\beta_{sd} - \beta_{.d})(Q_{tsd} - Q_{t.d}) + (\beta_{sd} - \beta_{.d})Q_{t.d} + \beta_{.d}(Q_{tsd} - Q_{t.d})\} \quad .
\end{aligned}
$$

Writing the conditional expectation of outcome for person i using district-level quality measures, we obtain

$$E(Y_{itsd} \mid Q_{t.d}) = \alpha_{.d} + \beta_{.d} Q_{t.d} + \text{Covariance } (\beta_{sd}, Q_{tsd} \mid Q_{t.d})$$

where the final term on the right-hand side is a conditional covariance. If districts with higher schooling expenditure are more utilitarian in the sense that Covariance ($\beta_{sd}, Q_{tsd} \mid Q_{t.d}$), increases

---

[32] The assumption of linearity is made for convenience. Strict linearity and utilitarianism defined over the aggregate of Y would allocate all funds to the district with the highest β_{sd} assuming that all schools have equal enrollment.

in $Q_{t,d}$, estimated effects of quality on earnings are upward stated as more aggregative schooling quality measures are used in micro regressions. They estimate an effect greater than the average for any school. This condition would be satisfied if, for example, all school districts were utilitarian and scale multiples of each other. (Proportional expansion of $Q_{t,d}$ would cause a proportional increase in $Q_{t,d}$ and would produce a positive relationship).

The distinction between this model and a measurement error model is as follows. If all schools are of identical size, and measurement error is independent and identically distributed, then any aggregate of N schools independent of governmental jurisdiction would have the same bias for β_d. If, however, the allocation model just presented is operative, estimated effects for any N schools should depend on whether schools are aggregated across governing units. For instance, in California where redistribution is egalitarian, aggregation across school districts should weaken any estimated schooling quality effect on earnings.

Thus, the two hypotheses to be considered and distinguished are:

H_0: <u>Measurement Error</u> -

Any N units (schools) when aggregated should show <u>diminished</u> bias. Let "^" denote the OLS estimator then

$$i.e. \quad plim\,\hat{\beta}_{.d} = \frac{\beta_{.d}}{1 + \frac{\sigma_\varepsilon^2}{N}} \; .$$

where σ_ε^2 is the variance of the measurement error. As N increases, plim $\hat{\beta}_d$ should decline for <u>any</u> collection of N schools, independently of whether aggregation is within or between jurisdictions. Here the number of blocks of N schools becomes large.

H_A: <u>Allocation Bias</u>

The aggregation effect is dependent on whether the aggregates are constructed within or between political jurisdictions. If Benthamite policies favor better districts, (i.e. there is a

positive dependence between the β and Q) the positive bias increases with higher levels of aggregation of schooling inputs across jurisdictions.

To distinguish between these two hypotheses, the following test could be performed. Estimate β_d within jurisdictions, then across jurisdictions. If the allocation bias model is operative, then for large N, cross-jurisdictional estimated effects of quality on earnings should increase. If the measurement error story is correct, aggregation within or between jurisdictions should have the same effect. If the redistributive story is correct, averaging within political units will have a different effect than averaging over political units. If the Benthamite story is correct at all levels of government, the estimated effect of quality on earnings increases with the level of aggregation. If the egalitarian story is correct the dependence should weaken with the level of aggregation.

VI. Summary and Conclusions

This paper examines the empirical and conceptual foundations of recent studies designed to estimate the impact of schooling quality on earnings. We set forth a general analysis of variance framework within which we nest previous models estimated in the literature. We make explicit the identifying assumptions used in the previous literature and the definitions of the quality effect utilized in previous studies. A variety of interesting economic models of the pricing of factor services emerge as special cases of the analysis of variance model. We also estimate more general models and test restrictions imposed in previous work.

We reach the following main conclusions about the quality of the evidence on schooling quality:

1. Methods based on aggregate measures of schooling quality used in micro regressions secure identification of quality effects by comparing the earnings of people born in the same state (and presumed to have been educated there) but living in different labor markets.

2. A crucial identifying assumption in the literature is the absence of selective migration on the basis of earnings in the region of residence.

58

3. However, nonparametric rank tests, evidence on the wages of migrants in destination regions, and evidence of strong region of birth - region of residence interactions all point to the empirical importance of selective migration. Under the assumption that everyone in the same state of birth receives the same quality of schooling, ranks of quality and earnings by region of birth should be correlated as well as invariant across regions of residence for each level of educational attainment. We find little relationship between the quality of schooling experienced by a cohort and the ranking of persons in regional wage distributions. Not only is the connection weak, but it changes across labor markets. Earnings ranks classified by state of birth show some stability over regions of residence, but the sources of the stability are not necessarily related to aggregate measures of quality.

4. Imposing their identifying assumptions, we find that Card and Krueger's analysis can be replicated for 1980 as well as for 1990 and 1970. Within their framework, the evidence of quality effects on earnings is especially strong in 1990. An interpretable macro regional model can be constructed within their framework.

5. The Card-Krueger assumption about the absence of region of birth-region of residence interactions is testable. It is rejected. When it is relaxed, estimated schooling quality effects weaken and sometimes become perverse in 1970 and 1980 data. The main results for 1990 hold up to the introduction of these interactions.

6. The assumption of the linearity of the earnings-schooling relationship is also tested and decisively rejected. In a nonlinear model, the only surviving evidence of any schooling quality effect is in the return to college education.

7. There is weak evidence that higher schooling quality raises schooling attendance. This confirms a finding in Card and Krueger and Johnson and Stafford.

8. Allowing for a more symmetric treatment of the regressors also weakens the quality-earnings connection. When the coefficients on experience and its square are permitted to vary over state of birth and region of residence, the quality effect all but disappears. There is little evidence that quality affects the rate of return to post-school work experience.

9. Introducing interactions between region of birth and region of residence in the second stage similarly dampens quality's effect on earnings.

10. Conventional efficiency unit models of the pricing of labor services do not receive support. A pricing model with heterogeneous human capital and regional shocks affecting less skilled labor receives considerable support.

Keats wrote that beauty is truth. On these grounds we favor the original Card and Krueger model. It holds up in 1970 and 1990 and leads to a richly interpretable economic model of the effect of quality on earnings. When extended to include aggregate demand and supply variables it produces consistent and sensible results. But Keats also wrote that truth is beauty. Taking as "truth" our statistical tests on nonlinearity and region of birth-region of residence interactions produces a less pretty but still economically interpretable description of the data that downplays the role of schooling quality in explaining earnings but emphasizes the role of heterogeneous human capital and regional labor markets.

Have We Settled the Debate?

The evidence in this paper, like the evidence in the literature that precedes it, is not decisive on the question of whether schooling inputs can raise earnings. All we have done is raise doubts about a certain type of aggregate evidence used by advocates of certain schooling policies.

The available measures of schooling quality are crude. The estimating equations used in the literature do not capture explicit choice mechanisms for parents and school authorities or the detailed schooling production processes required to justify specific policy interventions. As Hanushek (1991) notes, the absence of a strong empirical relationship between measured outcomes and measured inputs may reflect the quality of the measures and the inefficiency of most schooling organizations in transforming inputs to outputs.

60

References

Behrman, Jere R., and Birdsall, Nancy. "The Quality of Schooling: Quantity Alone Is Misleading." *A.E.R.*, 73 (December 1983): 928-46.

Ben-Porath, Yoram. "The Production of Human Capital and the Life Cycle of Earnings." *J.P.E.* 75, no.1 (August 1967): 352-65.

Betts, Julian. "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth." Unpublished paper, University of California, San Diego (1993).

_____, "Is There a Link Between School Inputs and Earnings? Fresh Scrutiny of an Old Literature," forthcoming in G. Burtless, ed. *Does Money Matter? The Link Between Schools, Student Achievement and Adult Success*, Brookings Institution, Washington, D.C. (1995).

Card, David and Krueger, Alan. "School Quality and Black-White Relative Earnings: A Direct Assessment." *Quarterly Journal of Economics*, 107 (February 1992a).

_____, "Does School Quality Matter: Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy*, 100, No. 1, (1992b): 1-40.

_____, "The Economic Return to School Quality: A Partial Survey." Unpublished working paper, Princeton University (1994).

Coleman, James S., et.al. *Equality of Educational Opportunity*. Washington: Government Printing Office (1966).

Hanushek, Eric. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature*, 24 (1986):1141-1177.

_____, "The Impact of Differential Expenditures on School Performance," *Educational Researcher*, 18 (1989): 45-65.

_____, "When School Finance 'Reform' May Not Be Good Policy", *Harvard Journal of Legislation*, Vol. 28 (June 1991): 423-464.

_____ et. al., "Aggregation Bias," Unpublished manuscript, University of Rochester, (1994).

_____ and Ralph W. Harbison. *Educational Performance of the Poor: Lessons from Rural Northeast Brazil*, Oxford University Press for the World Bank (1992).

Heckman, James, Anne Layne-Farrar and Petra Todd. "Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship," forthcoming in G. Burtless, ed. *Does Money Matter? The Link Between Schools, Student Achievement and Adult Success, Brookings Institution, Washington, DC (1995).*

_____ and Derek Neal. "James Coleman's Contributions to Education: An Intellectual Odyssey", forthcoming in *James S. Coleman: Falmer Sociology Series,* London-New York-Philadelphia, Falmer Press (1996).

_____ and Richard Robb. "Evaluating The Impact of Interventions on Outcomes", in J. Heckman and B. Singer, eds., *Longitudinal Analysis of Labor Market Data,* Cambridge, England, Cambridge University Press (1985).

_____ and Jose Scheinkman, "The Importance of Bundling in A Gorman-Lancaster Model of Earnings", Review of Economic Studies, (1987).

_____ and G. Sedlacek. "Heterogeneity, Aggregation and Market Wage Functions", *Journal of Political Economy* (December 1985): 1123-1148.

Hedges, Larry V., Richard D. Laine and Rob Greenwald. "Does Money Matter? A Meta-analysis of Studies of the Effects of Differential School Inputs on Student Outcomes." *Educational Researcher,* 23, No. 3 (April 1994): 514.

Johnson, George E., and Frank P. Stafford. "Social Returns to Quantity and Quality of Schooling." *Journal of Human Resources,* 8 (Spring 1973): 139-55.

Mincer, Jacob, "Investment in Human Capital and Personal Income Distribution", Journal of Political Economy, Vol. LXVI, 281-302, (August 1958).

_____, *Schooling, Experience and Earnings,* New York, National Bureau of Economic Research, 1974.

Morgan, James and Ismail Sirageldin. "A Note on the Quality Dimension in Education." *Journal of Political Economy,* 76 (September/October 1968): 1069-1077.

Rizzuto, Ronald, and Paul Wachtel. "Further Evidence on the Returns to School Quality", *Journal of Human Resources,* 15 (Spring 1980): 240-54.

Sattinger, Michael. *Capital and The Distribution of Labor Earnings,* North Holland (1980).

Solon, Gary and Thomas Hungerford. "Sheepskin Effects In The Returns To Education", *Review Economics and Statis*tics, 69 (February 1987): 175-177.

Speakman, Robert and Finis Welch. "Micro and Macro Estimates of The Value of School Resources", unpublished manuscript, Texas A.& M (June, 1994).

Tinbergen, Jan. *Income Distribution: Analysis and Policies*: Amsterdam (1975).

_____ ,"Income Distribution: Second Thoughts", *De Economist*, 125 (1977): 315-339.

Wachtel, Paul. "The Effects on Earnings of School and College Investment Expenditures. *"Review of Economics and Statistics*, 58 (August 1976): 326-31.

Welch, Finis. "Measurement of the Quality of Schooling." *American Economic Review Papers and Procedings*, 56 (May 1966): 379-92.

_____. "Labor-Market Discrimination: An Interpretation of Income Differences in the Rural South." *Journal of Political Economy* , 75 (June 1967): 225-40.

White, Halbert. "A Heteroskedasticity Consistent Covariance Matrix and a Direct Test for Heteroskedasticity," E*conometrica*, 48, (1980): 817-838.