The Dynamics of Educational Attainment for Blacks, Hispanics, and Whites

by

Stephen V. Cameron
Columbia University

and

James J. Heckman
University of Chicago and the American Bar Foundation

First Draft April, 1992
Revised, September 1998
Revised, November, 1999
Introduction

1. Our Data and Some Facts About Schooling Attendance

(a) Basic Facts about Schooling Attainment

(b) Dynamics of Schooling Transitions

(c) The Family Income-Schooling Connection
Explanatory Variables

(1) By age 15 before any significant dropping out occurs, there are large differences in racial-ethnic schooling attainment. Whites start ahead and stay ahead.

(2) Early schooling attainment is highly correlated with later schooling attainment.
(3) Most high school completion occurs at age 18 or 19 after 12 years of continuous schooling. Interruption of secondary schooling is an uncommon event. High school completion after age 19 is almost exclusively through GED attainment. College entry, if it happens at all, follows high school completion with little or no delay. About 90% of all college entry occurs within two years of high school completion.

(4) Family income is a powerful correlate of college entry. Indeed, differences in family income alone can explain
most of the White-minority
gaps in college entry of high school completers.

(5) Whites and minorities are significantly different in terms of family background, the actual costs of college they face, and a number of other important variables.
2. An Econometric Model of Schooling Attainment

Consider someone at age \( a \) facing choice set \( C_{a,j} \). \( C \in C_{a,j} \) is a choice taken. \( D_{a,j,c} \) is an indicator of the choice.

\[
\sum_{c \in C_{a,j}} D_{a,j,c} = 1.
\]

\[\hat{c}_{a,j} = \arg \max_{c \in C_{a,j}} \{ V_{a,j,c} \} \]

Linearize to obtain

(1) \[ V_{a,j,c} = Z'_{a,j} \beta_{a,j,c} + \varepsilon_{a,j,c} \]

where \( Z_{a,j} \) is a vector of ob-
\( \varepsilon_{a,j,c} = \alpha_{a,j,c} \eta + \nu_{a,j,c} \)

where

\( \eta \perp \perp \nu_{a,j,c} \)

(“\( \perp \perp \)” denotes independence)

for all \( a, j, c \), and the \( \eta \) are independent across persons, and \( \eta \) is a mean zero, unit variance random variable.
Would like to condition on $\eta$. This would eliminate person-specific unobservables.
(A-2) $\nu_{a,j,c}$ is an extreme value random variable, independent of all other $\nu_{a',j'',c''}$ except for $a = a', j = j''$ and $c = c'''$. 
(3) $\Pr(D_{a,j,c'} = 1 \mid Z_{a,j,\eta})$

$$= \Pr(\arg\max_c V_{a,j,c} = c' \mid Z_{a,j,\eta})$$

$$= \exp\left\{Z_{a,j} \beta_{a,j,c'} + \alpha_{a,j,c'\eta}\right\}$$

$$\sum_{c \in C_{a,j}} \exp\left\{Z'_{a,j} \beta_{a,j,c} + \alpha_{a,j,c\eta}\right\}.$$
nonlinear model is to integrate it out.

(A-3) \( Z_{a,j} \perp \perp \eta, \ \forall a, j \in C_{a,j} \)

for all choice sets
(4) \( \Pr(D_{a,c'} = 1 \mid Z_a, \eta) \)

\[
\exp\left\{ Z'_{a,c'} \beta_{a,c'} + \alpha_{a,c'} \eta \right\} \\
\sum_{c \in C_a} \exp\left\{ Z'_{a,c} \beta_{a,c} + \alpha_{a,c} \eta \right\}
\]

To form the probability of any life cycle string of choices, we can write
(5) \[ \text{Pr}(D_{a,c'} = 1 \mid Z_a, \eta) \cdot \]
\[ \text{Pr}(D_{a+1,c',c} = 1 \mid Z_{a+1,c',\eta}) \cdot \]
\[ \text{Pr}(D_{a+2,c_{a+1},c} = 1 \mid Z_{a+2,c_{a+1},\eta}) \]
\[ \cdots \text{Pr}(D_{\bar{a},c_{\bar{a}-1},c} = 1 \mid Z_{\bar{a},c_{\bar{a}-1},\eta}). \]
Models Extends the Literature

(A-4a) \[ C_{a,j} = C_j, \quad \forall a \]
(No age effects)

(A-4b) \[ C_j = \{j, j + 1\} \]
(No age effects pure grade transition model)
\( (A-4c) \quad \alpha_{a,j,c} = \alpha_{j,c} = 0 \)

\[ \forall a, j \]

(No\-heterogeneity bias)
3. Evidence on Educational Selectivity and the Dynamics of Schooling Choices: Estimates from the NLSY Data

Estimating The Baseline Econometric Model

Our strategy is as follows.
(1) For transitions with relatively few (less than 30) observations, we only estimate the intercepts and not the slope parameters in $\beta_{a,j,c}$. Factor loadings for these parameters are set to zero.

(2) We test for the presence of initial age ($a$), initial state ($j$) and final destination ($c$) interactions in the slope coefficients (denoted $\beta_{a,j,c}$) and factor loadings ($\alpha_{a,j,c}$).
(3) We test for differences in the slope coefficients and factor loadings among Blacks, Whites and Hispanics for the transitions that are not “rare” as defined in (1), maintaining ethnic/race-specific intercepts for each transition.

(1) Initial Grade Level.

(2) Secondary School Transitions for School Attenders.
(3) High School Dropouts.
(4) College Entry.
Goodness-of-Fit and the Importance of Unobserved Heterogeneity

Are There Differences in Racial-Ethnic Schooling Behavior?
Various Indices

(6a) \( E_{Z_w} \Pr(Z_w \hat{\beta}_w) - E_{Z_b} \Pr(Z_b \hat{\beta}_b) \)

\[ = E_{Z_w} \Pr(Z_w \hat{\beta}_w) - E_{Z_b} \Pr(Z_b \hat{\beta}_w) \]

\[ + E_{Z_b} (\Pr(Z_b \hat{\beta}_w) - \Pr(Z_b \hat{\beta}_b)) \]

\[ = \text{gap due to endowment difference} + \text{gap due to behavioral difference}. \]
\[(6b) \quad E_{Z_w}(Pr(Z_w \hat{\beta}_w) - Pr(Z_w \hat{\beta}_b))
\]

\[+ (E_{Z_w} Pr(Z_w \hat{\beta}_b) - E_{Z_b} Pr(Z_b \hat{\beta}_b))\]

= gap due to behavioral difference + gap due to endowment difference
\[(7a) \quad E_{Z_w}(\text{Pr}(Z_w \hat{\beta}_w) - \text{Pr}(Z_w \hat{\beta}_b)) \]

and

\[(7b) \quad E_{Z_b}(\text{Pr}(Z_b \hat{\beta}_w) - \text{Pr}(Z_b \hat{\beta}_b)). \]
Effects of Individual Variables

The Effects of College Tuition

Supporting Evidence From Other Studies

4. Summary and Conclusions