V. J. Hotz
June 3, 2007

Notes on Dahl, "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets"

Stylized fact: Rates of return to Schooling (College Degree relative to High School Degree) differ across regions of country (states).

Why? Why don't we get factor price equalization?

Could be different production functions with immobile factors which affect returns to labor, especially skills.

Could be differences in amenities across regions that draw different people to different regions.

If Comparative Advantage and/or amenity attraction in play, what is consequence for <u>observed</u> returns to college?

Selection bias in estimated returns &, thus, need to correct for selection.

Contributions of Paper:

Develops method for correcting for selectivity bias in state-specific returns to college degree (relative to high school) that deals with large number of markets (states)

Tests for implications of comparative advantage affecting mobility of workers by education

Model of Mobility & Earnings

Population earnings function for individual *I* if works in state *k*:

(1) $$y_{ik} = \alpha_k + x_i'\delta_k + s_i\beta_k + u_{ik}, \quad k = 1,\ldots,N,$$

Mobility decision based on utility maximization: For individual *i* born in state *j*, utility associated with move to state *k* given by:

(2) $$V_{ijk} = y_{ik} + t_{ijk}$$

where t_{ijk} is vector indexing tastes for moving from state *j* to state *k*.

The deviation of an individual's earnings if they were to work in state k from the average for the entire population (including individuals who do not actually work in state k) is

$$y_{ik} - E[y_{ik} | x_i, s_i] = u_{ik} \qquad (k = 1, \ldots, N). \tag{3}$$

Define a similar equation for the deviation of an individual's taste for moving from state j to state k from the population average, so that

$$t_{ijk} - E[t_{ijk} | z_i] = w_{ijk} \qquad (k = 1, \ldots, N), \tag{4}$$

where z_i is a vector of individual characteristics and w_{ijk} is an error term for individual deviations from mean tastes. Notice that I allow the value for mean tastes to be a function of both state of birth j and state of residence k, whereas I restrict mean earnings in (3) to be a function only of state of residence. Tastes for moving from state j to state k potentially include an overwhelming number of variables. For example, t_{ijk} could include the costs of moving from j to k, the difference in climate between j and k, the difference in state tax rates between j and k, or any other nonwage differences between the two states.

The expression for V_{ijk} can now be written in terms of the population mean and an error component specific to the individual:

$$V_{ijk} = V_{jk} + e_{ijk} \qquad (k = 1, \ldots, N), \tag{5}$$

where $V_{jk} = E[y_{ik} | x_i, s_i] + E[t_{ijk} | z_i]$ and $e_{ijk} = u_{ik} + w_{ijk}$. In the selection literature V_{jk} is often called the subutility function.

Individuals follow the migration path that maximizes their utility, so that individual i chooses to move from state j to state k according to

$$\begin{aligned} M_{ijk} &= 1 \quad \text{if and only if} \quad V_{ijk} = \max(V_{ij1}, \ldots, V_{ijN}), \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

where M_{ijk} is an indicator for whether individual i actually moves from state j to state k. The selection equations can alternatively be written as

$$\begin{aligned} M_{ijk} &= 1 \quad \text{if and only if} \quad V_{jk} + e_{ijk} \geq V_{jm} + e_{ijm} \quad \forall m, \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{6}$$

Utility depends on the specific migration path j to k; that is, the utility for an individual depends not only on the state of residence, but also on the state of birth. Assume the set $\{V_{ij1}, \ldots, V_{ijN}\}$ has a unique maximum and the error terms from the N selection criteria in (6) have a joint distribution that has finite moments and depends on a finite dimensional parameter set.

In this model, an individual can only live and work in one state; therefore, earnings for an individual are not observed in every state. The selection rule is

$$y_{ik} \text{ observed} \quad \text{if and only if} \quad M_{ijk} = 1, \tag{7}$$

taneously. Equations (1)–(7) describe an extended Roy model of earnings and mobility. Note that individuals currently living in state k are not a random sample of the population, and in general

$$
\begin{aligned}
(8) \qquad E[u_{ik}|y_{ik} \text{ observed}] &= E[u_{ik}|M_{ijk}=1] \\
&= E[u_{ik}|e_{ijm}-e_{ijk} \leq V_{jk}-V_{jm}, \forall m] \\
&\neq 0.
\end{aligned}
$$

Challenges of Estimating Roy Model with Many Sectors

Challenge 1: Control functions become difficult to formulate, as need different control function for each "birth state" j and all possible "residence states" i. Number of states here is 51 (50 plus District of Columbia)!

Dahl exploits and extends approximation strategy of Lee (1983, *Econometrica*) to reduce "curse of dimensionality" from 50 alternatives to 1.

Challenge 2: Have to model "tastes" for different localities that enter into the V_{ik}'s in (5).

Dahl shows can substitute locational choice probabilities for differences in V_{ik}'s. [Same idea used in Hotz and Miller, 1993, *REStud*, in single-agent DP structural choice models.]

3.1.1. *Using the Lee Approach to Selection Correction*

Accounting for the correlation of the error terms from N selection equations with the error term in the earnings equation of interest appears overwhelming. A parametric generalization of Heckman's two-step approach would require a complete specification of $f_{jk}(u_{ik}, e_{ij1} - e_{ijk}, \ldots, e_{ijN} - e_{ijk})$, and would involve the integration of an $(N-1)$-fold integral. Lee (1983) suggests reducing the dimensionality of the problem by reframing the N selection equations in (6) in terms of order statistics. Combining equations (6) and (7), the selection rule for state k becomes

$$y_{ik} \text{ observed} \quad \text{if and only if}$$
$$(V_{j1} - V_{jk} + e_{ij1} - e_{ijk}, \ldots, V_{jN} - V_{jk} + e_{ijN} - e_{ijk})' \leq \mathbf{0}$$

where $\mathbf{0}$ is an N-dimensional column vector. To understand Lee's approach, note that an equivalent expression is

$$(9) \qquad y_{ik} \text{ observed} \quad \text{if and only if} \quad \max_m (V_{jm} - V_{jk} + e_{ijm} - e_{ijk}) \leq 0$$

where $\max_m(\bullet)$ indicates the maximum over m. Thus any selectivity bias in y_{ik} is driven by the event that the maximum of the collection of random variables $V_{j1} - V_{jk} + e_{ij1} - e_{ijk}, \ldots, V_{jN} - V_{jk} + e_{ijN} - e_{ijk}$ is less than or equal to zero. The distribution function H_{jk} of the maximum order statistic, conditional on the subutility function differences, can be expressed as

$$(10) \qquad \begin{aligned} &H_{jk}(t|V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}) \\ &= Pr\Big[\max_m(V_{jm} - V_{jk} + e_{ijm} - e_{ijk}) < t|V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}\Big] \\ &= Pr[e_{ij1} - e_{ijk} < V_{j1} - V_{jk} + t, \ldots, e_{ijN} - e_{ijk} < V_{jN} - V_{jk} + t] \\ &= F^e_{jk}(V_{j1} - V_{jk} + t, \ldots, V_{jN} - V_{jk} + t), \end{aligned}$$

which makes clear that H_{jk} (conditional on $V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}$) evaluated at zero is simply the probability of sample selection. Given the equivalent formulation of the selection rule in (9), the cumulative distribution function F_{jk} can now be expressed in the following ways:

$$(11) \qquad \begin{aligned} &F_{jk}(r, V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}) \\ &= Pr(u_{ik} < r, e_{ij1} - e_{ijk} < V_{j1} - V_{jk}, \ldots, e_{ijk} - e_{ijN} < V_{jN} - V_{jk}) \\ &= Pr\Big[u_{ik} < r, \max_m(V_{jm} - V_{jk} + e_{ijm} - e_{ijk}) < 0|V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}\Big] \\ &= G_{jk}(r, 0|V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}) \end{aligned}$$

where G_{jk} is a well-defined cumulative joint distribution function for u_{ik} and $\max_m(V_{jm} - V_{jk} + e_{ijm} - e_{ijk})$. Writing (11) in terms of density functions provides another way to express this distributional equivalence:

$$(12) \qquad f_{jk}(u_{ik}, e_{ij1} - e_{ijk}, \ldots, e_{ijN} - e_{ijk} | V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk})$$
$$= g_{jk}\left(u_{ik}, \max_m(V_{jm} - V_{jk} + e_{ijm} - e_{ijk}) | V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}\right)$$

where both sides of the equation are explicitly written as conditional densities to emphasize the dependence on the differences in subutility functions. Equation (12) has reduced the dimensionality of the error terms that must be accounted for by expressing an N-variate joint distribution in terms of a bivariate distribution.

Key Reinterpretation of Lee's Approach

Lee's parameterization of distribution function in (12) is that u_{ik} and maximum order statistic *do not depend on differences in* $V_{jk'} - V_{jk}$, for all j, k'.

$$(A\text{-}1) \qquad g_k\left(u_{ik}, \max_m(V_{jm} - V_{jk} + e_{ijm} - e_{ijk})\right)$$

does not depend on $V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}$.

3.2.1. *Formulation as a Single-Index Model*

The formulation of mobility and earnings in equations (1) and (6) implies the earnings equations can be rewritten as multiple-index, partially-linear models:

$$(13) \qquad y_{ik} = \alpha_k + x_i'\delta_k + s_i\beta_k$$
$$+ \sum_{j=1}^{N}[M_{ijk} \times \psi_{jk}(V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk})] + \upsilon_{ik} \quad (k = 1, \ldots, N)$$

where $\psi_{ij}(\cdot) = E(u_{jk} | V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk})$ and υ_{ik} mean zero error term.

To take advantage of Lee's insight in a semiparametric framework, I make the following index sufficiency assumption:

$$\text{(A-2)}\qquad g_{jk}\Big(u_{ik}, \max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})|V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}\Big)$$
$$= g_{ik}\Big(u_{ik}, \max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})|p_{ijk}\Big)$$

where p_{ijk} is the probability that individual i moves from state j to state k given the vector $V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}$. The equivalence in (A-2) assumes that $p_{ijk}=p_{ijk}(V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk})$ exhausts all the information about how $V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}$ influences the joint distribution of u_{ik} and $\max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})$ contained in the sample. That is, the conditional distribution of u_{ik} and $\max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})$ can depend on the conditioning variables only through the single index p_{ijk}.

The single index p_{ijk} is the probability of an individual's first-best migration choice, a choice which is observable since the researcher knows where an individual chooses to live and work. This scalar migration probability associated with the maximum order statistic can be written in the following ways:

$$\text{(14)}\qquad \begin{aligned} p_{ijk} &= Pr(M_{ijk}=1|V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}) \\ &= Pr(V_{jk}+e_{ijk}\geq V_{jm}+e_{ijm},\ \forall m) \\ &= F^{e}_{jk}(V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}) \\ &= H_{jk}(0|V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}). \end{aligned}$$

The researcher must somehow account for the subutility functions to get an estimate of p_{ijk}, since the vector $V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}$ determines an individual's migration choice. Discussion of how to estimate p_{ijk} is postponed until later; for the moment, assume that an appropriate estimator is available.

Using assumption (A-2), equation (12) can be simplified to

$$\text{(15)}\qquad f_{jk}(u_{ik}, e_{ij1}-e_{ijk},\ldots,e_{ijN}-e_{ijk}|V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk})$$
$$= g_{jk}\Big(u_{ik}, \max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})|p_{ijk}\Big)$$

and the earnings equations can now be written as single-index, partially linear models:

$$\text{(16)}\qquad y_{ik}=\alpha_k+x_i'\delta_k+s_i\beta_k+\sum_{j=1}^{N}\{M_{ijk}\times\lambda_{jk}(p_{ijk})\}+\omega_{ik}\qquad (k=1,\ldots,N)$$

where for each birth state j, $\lambda_{jk}(\bullet)$ is an unknown function of the single index p_{ijk} and ω_{ik} is an error term. I refer to the λ_{jk}'s as the selection correction functions for state k. By construction, the error term ω_{ik} has zero mean given the migration probability and the fact that earnings are observed in a state:

$$E[\omega_{ik}|x_i, s_i, p_{ijk}, M_{ijk}=1]=0 \qquad (k=1,\ldots,N).$$

A proof for the result that $\psi_{jk}(V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk})=\lambda_{jk}(p_{ijk})$ if assumption (A-2) holds is provided in Appendix B.

3.2.2. *Extension to a Multiple-Index Framework*

One interpretation for the use of p_{ijk} in the correction function for polychotomous choice models relies on the fact that, subject to an invertibility condition[10]

$$\text{(17)} \qquad g_{jk}\Big(u_{ik}, \max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})|V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}\Big) = g_{jk}\Big(u_{ik}, \max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})|p_{ij1},\ldots,p_{ijN}\Big),$$

which simply states that the multiple migration probabilities, $p_{ij1},\ldots,p_{ijN}$, contain the same information as the differenced subutility functions, $V_{j1}-V_{jk},\ldots,$ $V_{jN}-V_{jk}$. This implies the earnings equations can be rewritten as multiple-index, partially-linear models that depend on all N migration probabilities:

$$\text{(18)} \qquad y_{ik}=\alpha_k+x_i'\delta_k+s_i\beta_k + \sum_{j=1}^{N}[M_{ijk}\times\mu_{jk}(p_{ij1},\ldots,p_{ijN})]+v_{ik} \qquad (k=1,\ldots,N)$$

where $\mu_{jk}(\bullet) = E[u_{ik}|p_{ij1},\ldots,p_{ijN}] = E[u_{ik}|V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}]$. Assumption (A-2) simplifies this equivalence by assuming that only the probability of the utility maximizing choice matters for the parameterization of the joint distribution g_{jk}. Hence, (A-2) can also be thought of as an exclusion restriction in that it requires the distribution of u_{ik} and $\max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})$ given $p_{ij1},\ldots,p_{ijN}$ to be the same as that given p_{ijk}.

A relaxation of (A-2) allows other probabilities besides the first-best choice probability to also influence the joint distribution g_{jk}. Letting $\vec{q}$ represent a chosen subset of the full set of migration probabilities $\{p_{ij1},\ldots,p_{ijN}\}$, this less restrictive assumption can be written as

$$\text{(A-3)} \qquad g_{jk}\Big(u_{ik}, \max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})|V_{j1}-V_{jk},\ldots,V_{jN}-V_{jk}\Big) = g_{jk}\Big(u_{ik}, \max_m(V_{jm}-V_{jk}+e_{ijm}-e_{ijk})|p_{ijk},\vec{q}\Big).$$

[10] To insure that equation (17) holds locally, the assumptions of the implicit function theorem must be satisfied. The $N\times N$ determinant of the vector of implicit equations $[F_{jm}^{e}(V_{j1}-V_{jm},\ldots,V_{jN}-V_{jm})-p_{ijm}]=0$, $m=1,\ldots,N$, must be nonzero so that the Jacobian is nonzero and a local inverse function exists.

In the current application, I end up adding the retention probability as another term in the correction functions. I discuss how this term was chosen for inclusion in the next section of the paper. Thus, the maintained distributional assumption for the current application is

$$\text{(A-4)}\qquad g_{jk}\Big(u_{ik}, \max_m(V_{jm} - V_{jk} + e_{ijm} - e_{ijk}) | V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk}\Big)$$
$$= g_{jk}\Big(u_{ik}, \max_m(V_{jm} - V_{jk} + e_{ijm} - e_{ijk}) | p_{ijk}, p_{ijj}\Big),$$

which implies the earnings equations can be written as

$$\text{(19)}\qquad y_{ik} = \alpha_k + x_i'\delta_k + s_i\beta_k + \sum_{j=1}^{N}\{M_{ijk} \times \lambda^*_{jk}(p_{ijk}, p_{ijj})\} + \omega^*_{ik} \qquad (k = 1, \ldots, N).$$

The correction terms in the wage equations for movers are now unknown functions of two probabilities, p_{ijj} and p_{ijk}. Notice that for stayers the correction terms are a function of a single probability, p_{ikk}, since $j = k$ for individuals who do not move from their birth state.

Index Sufficiency Assumption in (A-2)

Since the selection correction functions for a state depend on the joint distribution of u_{ik} and $e_{ij1} - e_{ijk}, \ldots, e_{ijN} - e_{ijk}$, it is informative to consider the bivariate covariances between the error term in the earnings equation and the error term in each of the migration equations:

$$\text{(20)}\qquad \text{cov}(u_{ik}, e_{ij1} - e_{ijk}), \ldots, \text{cov}(u_{ik}, e_{ijN} - e_{ijk}).$$

Each bivariate covariance in (20) can be broken up into four separate covariances. For example, the first term can be expressed as

$$\text{cov}(u_{ik}, e_{ij1} - e_{ijk}) = \text{cov}(u_{ik}, u_{i1}) - \text{var}(u_{ik}) + \text{cov}(u_{ik}, w_{ij1})$$
$$- \text{cov}(u_{ik}, w_{ijk}).$$

One can show that (A-2) is met if assume:

$$\text{(21)}\qquad u_{ik} = a_i + b_{ik}, \quad k = 1, \ldots, N$$

where a_i is individual fixed effect and b_{ik} is state-specific error term. (See paper.)

Now, however, that slight modification of (21)

$$\text{(22)}\qquad u_{ik} = \tau_k a_i + b_{ik}, \quad k = 1, \ldots, N$$

does not meet (A-2).

Implementation of Model for Estimation:

Dahl makes further simplifying assumptions to reduce "dimensionality" of control functions. (see paper)

Location choice probabilities estimated for "cells" of individuals with identical characteristics:

$$(24) \qquad p_{ijk} = Pr(M_{ijk} = 1 | V_{j1} - V_{jk}, \ldots, V_{jN} - V_{jk})$$
$$= Pr(M_{ijk} = 1 | \text{cell}).$$

Use series estimator to approximate $\lambda^*_{jk}, j,k = 1,\ldots,N$ in control functions for (19):

$$(25) \qquad \lambda^*_k(p_{ijk}, p_{ijj}) \cong \sum_{t=1}^{T} \kappa^t_k b^t_k(p_{ijk}, p_{ijj})$$

where the functions $b^t_k(\bullet)$ are referred to as the basis functions. Similar approximations exist for the stayers' correction functions. Two common choices for basis functions are the terms of a polynomial or Fourier series. Since both choices yield

Estimation Results

TABLE I

SUMMARY STATISTICS

Variable	U.S.	California	Florida	Illinois	Kansas	New York	Texas
(1) Migrant (%)	31	—	—	—	—	—	—
	(0.1)						
(2) Inmigrant (%)	—	37	69	20	33	12	33
		(0.2)	(0.3)	(0.2)	(0.6)	(0.2)	(0.2)
(3) Outmigrant (%)	—	25	34	31	41	32	22
		(0.2)	(0.4)	(0.3)	(0.6)	(0.2)	(0.2)
(4) Less than High School (%)	13	11	15	9	10	10	15
	(0.1)	(0.1)	(0.2)	(0.2)	(0.4)	(0.2)	(0.2)
(5) High School (%)	38	28	35	37	40	35	33
	(0.1)	(0.2)	(0.3)	(0.3)	(0.6)	(0.2)	(0.2)
(6) Some College (%)	28	35	30	30	31	29	30
	(0.1)	(0.2)	(0.3)	(0.3)	(0.6)	(0.2)	(0.2)
(7) College Graduate (%)	17	19	16	19	16	20	18
	(0.1)	(0.2)	(0.2)	(0.2)	(0.5)	(0.2)	(0.2)
(8) Advanced Degree (%)	4	6	4	6	3	7	5
	(0.0)	(0.1)	(0.1)	(0.1)	(0.2)	(0.1)	(0.1)
(9) Married (%)	63	54	59	63	68	57	68
	(0.1)	(0.2)	(0.3)	(0.3)	(0.6)	(0.3)	(0.2)
(10) Residence in SMSA (%)	64	95	83	70	32	74	72
	(0.1)	(0.1)	(0.2)	(0.3)	(0.6)	(0.2)	(0.2)
(11) Wage	11.93	14.38	11.15	12.78	10.29	13.72	11.27
	(0.01)	(0.04)	(0.05)	(0.05)	(0.07)	(0.05)	(0.04)
(12) Observations	538,953	51,150	24,316	26,792	6,045	38,139	37,846

Note: Standard errors in parentheses.
Source: 1990 U.S. Census data for white males, age 25–34, and employed full-time.

TABLE II

SUMMARY OF THE CELL MIGRATION PROBABILITIES

Education	Number of Cells[a]	Mean	Std. Dev.	10th Percentile	90th Percentile
		STAYERS			
Less than High School	616	0.6972	0.1243	0.5417	0.8361
High School Graduate	692	0.6790	0.1422	0.4783	0.8287
Some College	668	0.5997	0.1523	0.4000	0.7686
College Graduate	561	0.5325	0.1626	0.3158	0.7381
Advanced Degree	343	0.4857	0.1668	0.2857	0.7143
		MOVERS			
Less than High School	3923	0.0172	0.0281	0.0018	0.0429
High School Graduate	6090	0.0107	0.0218	0.0009	0.0261
Some College	5879	0.0136	0.0250	0.0012	0.0339
College Graduate	5159	0.0182	0.0311	0.0018	0.0436
Advanced Degree	3048	0.0298	0.0406	0.0038	0.0698

[a]Cells with 10 or fewer observations are excluded.

TABLE III

ESTIMATED WAGE EQUATIONS FOR CALIFORNIA, FLORIDA, ILLINOIS, KANSAS, NEW YORK, AND TEXAS

	California		Florida		Illinois	
	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected
(1) Less than High School	−0.1597	−0.1489	−0.1527	−0.1520	−0.1710	−0.1898
	(0.0082)	(0.0082)	(0.0101)	(0.0105)	(0.0113)	(0.0116)
(2) Some College	0.1383	0.1505	0.1337	0.1041	0.1165	0.0968
	(0.0059)	(0.0061)	(0.0080)	(0.0087)	(0.0076)	(0.0079)
(3) College Graduate	0.4378	0.4313	0.4485	0.4022	0.3645	0.3272
	(0.0075)	(0.0079)	(0.0106)	(0.0127)	(0.0100)	(0.0117)
(4) Advanced Degree	0.5996	0.5760	0.6407	0.5880	0.5461	0.5059
	(0.0110)	(0.0117)	(0.0172)	(0.0194)	(0.0147)	(0.0178)
(5) Experience	0.0778	0.0745	0.0663	0.0649	0.0580	0.0525
	(0.0075)	(0.0075)	(0.0107)	(0.0107)	(0.0097)	(0.0097)
(6) Experience Squared	−0.0023	−0.0023	−0.0024	−0.0023	−0.0008	−0.0005
	(0.0007)	(0.0007)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
(7) Experience	−0.0001	0.0001	0.0018	0.0017	−0.0034	−0.0041
Cubed × 100	(0.0018)	(0.0018)	(0.0026)	(0.0025)	(0.0026)	(0.0026)
(8) Married	0.1906	0.1438	0.1763	0.1714	0.1925	0.1736
	(0.0047)	(0.0056)	(0.0065)	(0.0070)	(0.0063)	(0.0069)
(9) Residence in SMSA	0.1754	0.1834	0.1146	0.1160	0.2496	0.2521
	(0.0109)	(0.0109)	(0.0084)	(0.0085)	(0.0067)	(0.0067)
(10) Wald test for λ	—	88.29	—	87.72	—	49.75
(Movers Only)		[0.0000]		[0.0000]		[0.000]
(11) Wald test for λ	—	1563.56	—	22.23	—	61.65
(Stayers Only)		[0.0000]		[0.0000]		[0.0000]
(12) Wald test for λ	—	463.56	—	117.57	—	109.96
		[0.0000]		[0.0000]		[0.0000]
(13) R-squared	0.1534	0.1606	0.1624	0.1668	0.1856	0.1891
(14) Observations	51,149	51,149	24,315	24,315	26,791	26,791

	Kansas		New York		Texas	
	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected
(1) Less than High School	−0.1887	−0.1933	−0.1958	−0.1985	−0.2023	−0.2046
	(0.0230)	(0.0233)	(0.0099)	(0.0099)	(0.0085)	(0.0085)
(2) Some College	0.0468	0.0349	0.1521	0.1297	0.1614	0.1356
	(0.0153)	(0.0165)	(0.0068)	(0.0074)	(0.0068)	(0.0071)
(3) College Graduate	0.3213	0.2863	0.4310	0.3977	0.5184	0.4697
	(0.0211)	(0.0253)	(0.0085)	(0.0107)	(0.0087)	(0.0098)
(4) Advanced Degree	0.4811	0.4122	0.5898	0.5495	0.6835	0.6130
	(0.0369)	(0.0462)	(0.0118)	(0.0145)	(0.0137)	(0.0153)
(5) Experience	0.0107	0.0110	0.0869	0.0820	0.0834	0.0811
	(0.0267)	(0.0268)	(0.0081)	(0.0081)	(0.0083)	(0.0083)
(6) Experience Squared	0.0028	−0.0028	−0.0041	−0.0038	−0.0029	−0.0027
	(0.0026)	(0.0026)	(0.0007)	(0.0007)	(0.0007)	(0.0007)

TABLE III—*Continued*

	Kansas		New York		Texas	
	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected
(8) Married	0.1790	0.1820	0.1881	0.1748	0.1901	0.2001
	(0.0134)	(0.0135)	(0.0055)	(0.0058)	(0.0057)	(0.0058)
(9) Residence in SMSA	0.2308	0.2296	0.2209	0.2225	0.1234	0.1139
	(0.0135)	(0.0138)	(0.0061)	(0.0061)	(0.0060)	(0.0060)
(10) Wald test for λ	—	7.37	—	85.08	—	43.57
(Movers Only)		[0.1946]		[0.0000]		[0.0000]
(11) Wald test for λ	—	2.31	—	60.67	—	116.81
(Stayers Only)		[0.3156]		[0.0000]		[0.0000]
(12) Wald test for λ	—	8.34	—	132.59	—	110.39
		[0.3037]		[0.0000]		[0.0000]
(13) R-squared	0.1574	0.1589	0.1912	0.1938	0.1932	0.1974
(14) Observations	6,044	6,044	38,138	38,138	37,845	37,845

Note: Standard errors in parentheses, p-values in brackets; both adjusted for the sampling variability of the estimated migration probabilities appearing in the correction functions (see footnote 24).

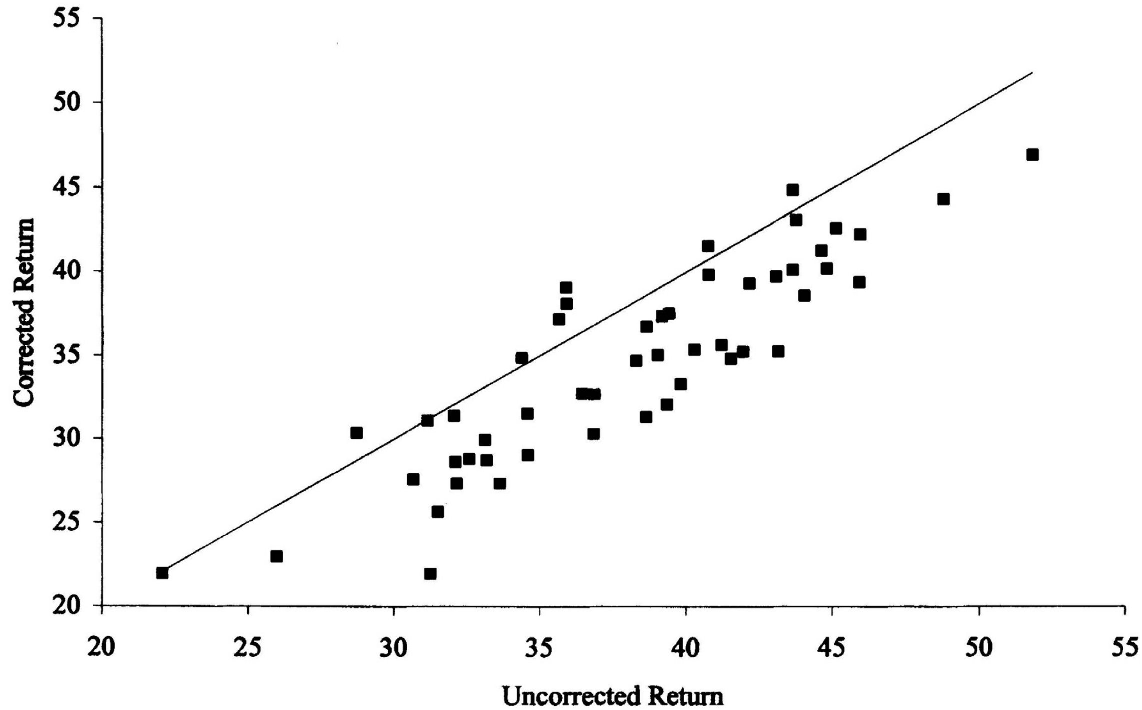

FIGURE 3.—Corrected versus uncorrected returns to a college education by state.

j to state k for college-educated movers in terms of earnings and amenities is

$$(26) \qquad \ln(p_{jk}^{CD}) = \theta_0^{CD} + \theta_1(y_k^{CD} - y_j^{CD}) + \theta_2^{CD}(A_k - A_j) + \theta_3^{CD} D_{jk} + \nu_{jk}^{CD}$$

where $\ln(\bullet)$ denotes the natural log, y_k^{CD} represents the average earnings of individuals with a college degree in state k, A_k is a vector of amenity variables associated with state k, D_{jk} is a vector of cost variables for moving from j to k, and ν_{jk}^{CD} is an error term. Note that for estimation, one would need to substitute in estimates of y_k^{CD} and y_j^{CD} since their true values are unavailable. Define a similar equation for individuals with a high school education, superscripting the appropriate variables and coefficients with an "HS" instead of a "CD". Equation (26) formalizes the assumption that migration flows are determined by earnings and amenity differences across states. Notice that schooling level does not change the package of amenities offered by a state. However, the value individuals in different education classes place on those amenities is expected to differ, which accounts for the education-specific coefficients on these variables. In contrast, while state-specific earnings depend on education level, the coefficient θ_1 is not superscripted by schooling level in the migration flow equations. The implication is that the log of college and high school migration flows respond identically to a given difference in earnings.

Differencing the log migration flows of college- and high school-educated individuals yields

$$(27) \qquad \begin{aligned} \ln(p_{jk}^{CD}) - \ln(p_{jk}^{HS}) = {} & (\theta_0^{CD} - \theta_0^{HS}) + \theta_1(y_k^{CD} - y_k^{HS}) \\ & - \theta_1(y_j^{CD} - y_j^{HS}) + (\theta_2^{CD} - \theta_2^{HS})(A_k - A_j) \\ & + (\theta_3^{CD} - \theta_3^{HS})D_{jk} + (\nu_{jk}^{CD} - \nu_{jk}^{HS}). \end{aligned}$$

Assuming the only component of earnings that differs by schooling level across states is the return to education, the expression $y_k^{CD} - y_k^{HS}$ represents the return to a college education relative to high school in state k. This relative return is simply the coefficient on the college dummy in the earnings equation, which I denote as β_k^{CD} for state k. Making this substitution and simplifying the notation of equation (27),

$$(28) \qquad \ln(p_{jk}^{CD}) - \ln(p_{jk}^{HS}) = \theta_0 + \theta_1 \Delta\beta^{CD} + \theta_2 \Delta A + \theta_3 D_{jk} + \nu_{jk}$$

where $\theta_0 = \theta_0^{CD} - \theta_0^{HS}$, $\theta_2 = \theta_2^{CD} - \theta_2^{HS}$, $\theta_3 = \theta_3^{CD} - \theta_3^{HS}$, $\Delta A = A_k - A_j$, $\nu_{jk} = \nu_{jk}^{CD} - \nu_{jk}^{HS}$, and $\Delta\beta^{CD} = \beta_k^{CD} - \beta_j^{CD}$.

Since the true value of $\Delta\beta^{CD}$ is not available, I substitute an estimate into equation (28) using results from Step 2. The coefficient estimates from a simple

TABLE V

Responsiveness of College Relative to High School Migration Flows to Differences in the Return to College and Amenities

Dependent Variable: $\ln(p_{jk}^{CD}) - \ln(p_{jk}^{HS})$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	0.51***	0.51***	0.54***	0.52***	0.52***	0.53***	0.52***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
Δ Corrected Return to College	2.29***	—	2.38***	1.89***	3.45***	2.39***	2.89***
	(0.4)		(0.35)	(0.39)	(0.62)	(0.40)	(0.77)
Δ Uncorrected Return to College	—	2.24***	—	—	—	—	—
		(0.32)					
Distance in Miles	—	—	−0.23	—	—	—	0.10
			(0.22)				(0.26)
Δ Unemployment Rate	—	—	−3.18**	—	—	—	−1.03
			(1.45)				(2.66)
Included Amenity Variables							
Quality of Life[a]				×			×
Climate[b]					×		×
State Spending and Taxing[c]						×	×
F test for Amenity Variables	—	—	—	10.83	3.92	7.70	5.44
				[.0000]	[.0000]	[.0000]	[.0000]
R-squared	0.0452	0.0479	0.0507	0.1325	0.0907	0.0855	0.1721
Observations	1,871	1,871	1,871	1,871	1,871	1,706	1,706

Notes: Huber-White standard errors in parentheses, *p*-values in brackets; the standard errors and *F* tests are adjusted for the sampling variability of the estimated state returns to a college education (see footnote 32). The symbol Δ represents the difference operator for the value of a variable between state *k* and state *j*. All explanatory variables are averages of 1980 and 1990 values except for the climate variables, which are already long-term averages. See Appendix D for variable sources and definitions.

[a] Quality of Life variables are Δ Population Density, Δ Doctors per Capita, Δ Dentists per Capita, Δ Hospital Costs, Δ Teacher's Salaries, Δ School Expenditures per Capita, Δ School Expenditures per Pupil, Δ Crime Rate, Δ Violent Crime Rate, and Δ Incarceration Rate.

[b] Climate variables are Δ Average Temperature, Δ Maximum Temperature, Δ Minimum Temperature, Δ Afternoon Humidity, Δ Annual Precipitation, Δ Number of Rainy Days, Δ Number of Sunny Days, and Δ Average Wind Speed.

[c] State Spending and Taxing variables are Δ State Spending on Education, Δ State Spending on Health and Human Services, Δ State Spending on Highways, Δ State Spending on Public Welfare, Δ Miscellaneous State Spending, Δ State Sales Tax, and Δ Average State Income Tax.

***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.