Spatial Variation in Higher Education Financing and the Supply of College Graduates

John Kennan*
University of Wisconsin-Madison and NBER

December 2017

Abstract

In the U.S. there are large differences across States in the extent to which college education is subsidized, and there are also large differences across States in the proportion of college graduates in the labor force. State subsidies are apparently motivated in part by the perceived benefits of having a more educated workforce. The paper extends the migration model of Kennan and Walker (2011) to analyze how geographical variation in college education subsidies affects the migration decisions of college graduates. The model is estimated using NLSY data, and used to quantify the sensitivity of migration and college enrollment decisions to differences in expected net lifetime income, focusing on how cross-State differences in public college financing affect the educational composition of the labor force. The main finding is that these differences have substantial effects on college enrollment, and that these effects are not dissipated through migration.

1 Introduction

There are substantial differences in subsidies for higher education across States in the U.S. Are these differences related to the proportion of college graduates in each State? If so, why? Do the subsidies change decisions about whether or where to go to college? If State subsidies induce more people to get college degrees, to what extent does this additional human capital tend to remain in the State that provided the subsidy?

There is a considerable amount of previous work on these issues, discussed in Section 3 below. What is distinctive in this paper is that migration is explicitly modeled. Recent work on migration has emphasized that migration involves a sequence of reversible decisions that respond to migration

*Department of Economics, University of Wisconsin, 1180 Observatory Drive, Madison, WI 53706; jkennan@ssc.wisc.edu. Earlier versions of this paper had a different title: "Higher Education Subsidies and Human Capital Mobility." I thank Gadi Barlevy, John Bound, Rebecca Diamond, Eric French, Chao Fu, Ahu Gemici, Grey Gordon, Joseph Han, Jim Heckman, Tom Holmes, Lisa Kahn, Maurizio Mazzocco, Bob Miller, Derek Neal, Jack Porter, Mike Rothschild, Jean-Marc Robin, Chris Taber, Sarah Turner, Jim Walker, Yoram Weiss and many seminar participants for helpful comments. I thank four referees and the Editor for detailed and constructive comments on both substance and exposition.

incentives in the face of potentially large migration costs.[1] The results of Kennan and Walker (2011) indicate that labor supply responds quite strongly to geographical wage differentials and location match effects, in a life-cycle model of expected income maximization.

This suggests that variation in tuition rates across States should have substantial effects on schooling decisions. This paper considers these effects in a dynamic programming model that allows for migration both before and after acquiring a college degree. In the absence of moving costs, the optimal policy for someone who decides to go to college is to move to the location that provides the cheapest education, and subsequently move to the labor market that pays the highest wage. At the other extreme, if moving costs are very high, the economic incentive to go to college depends only on the local wage premium for college graduates, and estimates based on the idea of a national labor market are likely to be misleading. Thus it is natural to consider college choices and migration jointly in a model that allows for geographical variation in both the costs and benefits of a college degree.

The model is estimated using panel data from the National Longitudinal Survey of Youth (1979 Cohort); the estimation sample includes white males with at least a high school education born between 1958 and 1964, observed annually from 1979 to 1994. College graduation is modeled as the outcome of a stochastic process, rather than as a choice: high school graduates choose whether to enroll in college, knowing that they may or may not emerge as college graduates, with probabilities that depend on their ability, and on the type of college. Enrollment and migration choices are affected by cross-State differences in college tuition and expenditures on higher education, and also by spatial wage differences. The main empirical finding is that cross-State differences in higher education expenditures, and (especially) in tuition levels, have substantial and long-lasting effects on the proportion of college-educated workers in each State.

2 Geographical Distribution of College Graduates

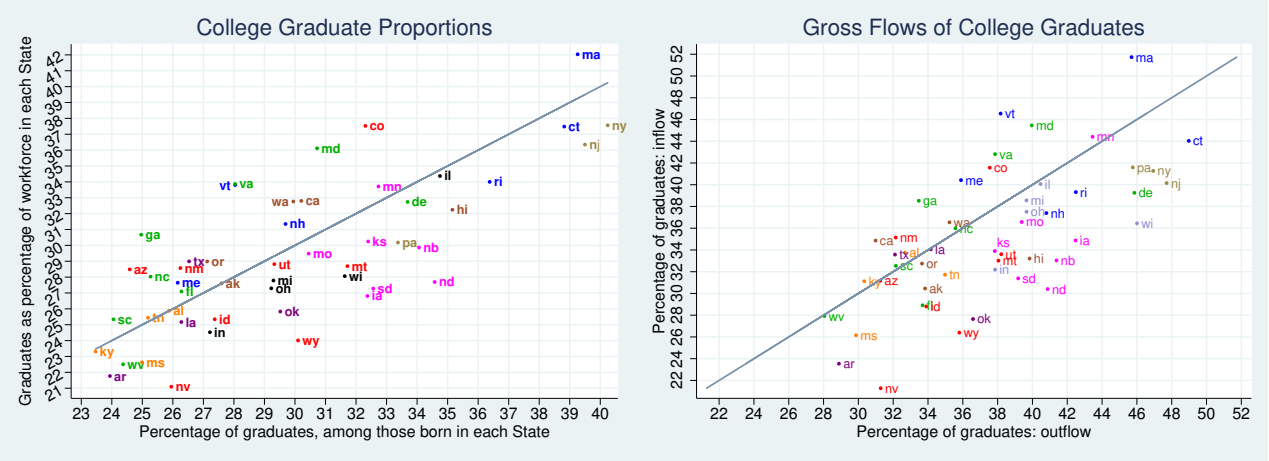
There are surprisingly big differences across States in the proportion of college graduates among those born in each State, and in the proportion of college graduates among those working in the State. Figure 2.1 shows the distribution of college graduates aged 25-50 in the 2000 Census, as a proportion of the number of people in this age group working in each State, and as a proportion of the number of workers in this age group who were born in each State. For example, someone who was born in New York is almost twice as likely to be a college graduate as someone born in Kentucky, and someone working in Massachusetts is twice as likely to be a college graduate as someone working in Nevada. Generally, the proportion of college graduates is high in the Northeast, and low in the South.[2]

There are also big differences in the proportion of college graduates who stay in the State where they were born. On average, about 45% of all college graduates aged 25-50 work in the State where they were born, but this figure is above 65% for Texas and California, and it is below 25% for Alaska

[1] See Kennan and Walker (2011),Gemici (2011) and Bishop (2012).

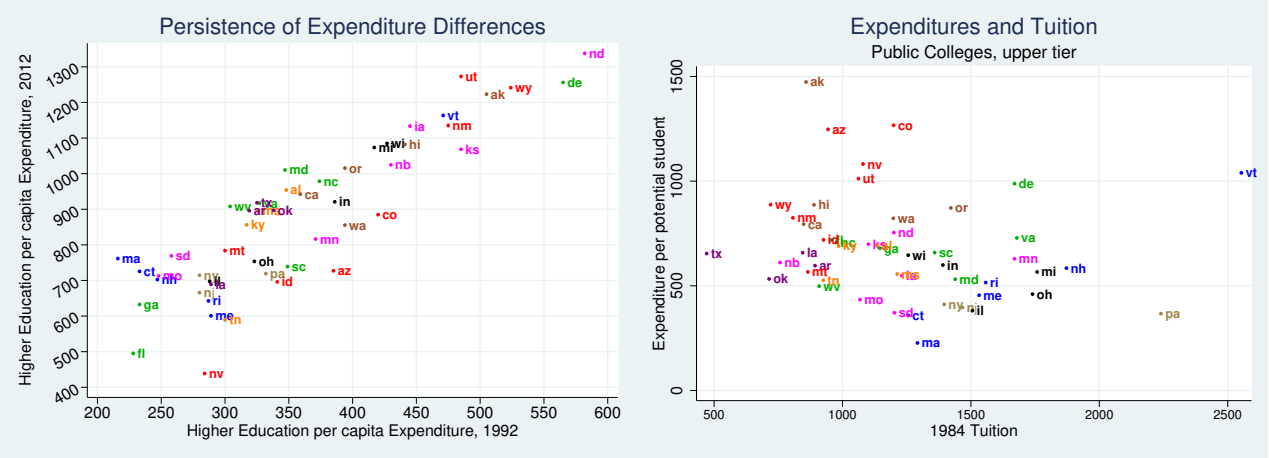
[2] The colors in Figure 2.1 (and subsequent figures) represent the nine Census Divisions.

Figure 2.1: Birth and Work Locations of College Graduates, 2000



and Wyoming. One might expect that the proportion of college graduates in the flow of in-migrants would be relatively high in States that have relatively few graduates in the native population, and similarly that the proportion of graduates in the flow of out-migrants would be high in States that have a high proportion of graduates in the native population. The right panel of Figure 2.1 shows that this is not the case.[34]

Figure 2.2: Tuition and Higher Education Expenditures



[3]Card and Krueger (1992) analyzed the effect of school quality using the earnings of men in the 1980 Census, classified according to when they were born, where they were born, and where they worked. An essential feature of this analysis is that the effect of school quality is identified by the presence in the data of people who were born in one State and who worked in another State (within regions, since the model allows for regional effects on the returns to education). This ignores the question of why some people moved and others did not.

[4]One possible explanation for the differences in the proportion of college graduates across States is that there are similar differences across States in the proportion of college graduates in the parents' generation, and there is a strong relationship between the education levels of parents and children. Of course this "explanation" merely shifts the question to the previous generation, but it is still of interest to know whether parental education is enough to account for most of the observed differences in college choices.There is indeed a strong relationship between the proportion of college graduates by State of birth for men aged 30-45 in the 2000 Census and the proportion of college graduates among the fathers in the 1970 Census, by State of residence (the regression coefficient is .78). But there is still plenty of inter-State variation in college graduation rates, even after controlling for the proportion of fathers who are college graduates.

2.1 Tuition and Higher Education Expenditures

States spend substantial amounts on higher education, and there are large and persistent differences in these expenditures across States. Figure 2.2 shows the variation in (nominal) per capita expenditures across States in 1992 and 2012, using data from the Census of Governments. The magnitude of these expenditures suggests that a more highly educated workforce is a major goal of State economic policies, perhaps because of human capital externalities. Figure 2.2 also shows tuition levels in public universities in 1984 (when the people in the NLSY cohort were aged 19 to 25), plotted against a measure of expenditures per student in these universities. Although differences in tuition levels and expenditures are correlated across States, there is considerable independent variation in these variables.

A common assumption in the literature on the relationship between college enrollment and cost is that the relevant measure of tuition is the in-state tuition level, given that most students attend college in their home State. This is a crude approximation. On average, about 20% of college freshmen in 2012 enrolled in an out of State college[5]. Moreover, this proportion varies greatly across States. At one extreme, the proportion of both imported and exported students was close to 10% for California and Texas.[6] At the other extreme, most of the freshmen in Vermont were not from Vermont, while most students from New Jersey were not studying in New Jersey.

3 Related Literature

The literature on the effects of State differences in college tuition levels is summarized by Kane (2006, 2007). The "consensus" view is that these effects are substantial – that a $1,000 reduction in tuition increases college enrollment by something like 5%.[7] Of course a major concern is that the variation in tuition levels across States is not randomly assigned, and there may well be important omitted

[5] See nces.ed.gov/programs/digest/d13/tables/dt13_309.20.asp?current=yes

[6] The proportion of imported students is the number of nonresident students as a fraction of total enrollments in the State, while the proportion of exported students is the number of students from this State attending college out of state, as a proportion of all students from this State.

[7] Keane and Wolpin (1997) used a dynamic programming model to analyze schooling and early career decisions in a national labor market; they estimated that a $2,000 tuition subsidy, at the national level, would increase college graduation rates by 8.4%. But this estimate was not based on observed responses to variation in actual tuition levels (indeed the data set did not contain any information about tuition). Instead, the model was used to infer the cost of attending college from data on wages of high school and college graduates, under the assumption that observed schooling decisions were made so as to maximize expected net present values of lifetime earnings. The essential argument can be illustrated by considering a hypothetical data set in which 50% of the people choose college, and the present value of lifetime earnings is $10,000 higher for college graduates, and there are zero-mean normally distributed payoff shocks attached to each choice. The estimated tuition cost would then be $10,000, and the effect of a $2,000 tuition subsidy would be inferred from the estimated distribution of the payoff shocks, by computing the probability that the shocks favor the non-college alternative by a positive amount less than $2,000.

Keane and Wolpin (2001) also estimated a related model of college choices, emphasizing the relationship between parental resources, borrowing constraints, and college enrollment (but with no consideration of spatial differences). A major result is that borrowing constraints are binding, and yet they have little influence on college choice. Instead, borrowing constraints affect consumption and work decisions while in college: if borrowing constraints were relaxed, the same people would choose to go to college, but they would work less and consume more while in school.

variables that are correlated with tuition levels.[8] There is no fully satisfactory way to deal with this problem. One approach is to use large changes in the net cost of going to college induced by interventions such as the introduction of the Georgia Hope Scholarship, as in Dynarski (2000), or the elimination of college subsidies for children of disabled or deceased parents, as in Dynarski (2003), or the introduction of the D.C. Tuition Assistance Grant program, as in Kane (2007). Broadly speaking, the results of these studies are not too different from the results of studies that use the cross-section variation of tuition levels over States, suggesting that the endogeneity of tuition levels might not be a major problem.[9] A detailed analysis of this issue would involve an analysis of the political economy of higher education subsidies in general, and of tuition levels in particular. For example, a change in the party controlling the State legislature or the governorship might be associated with a change in higher education policies, and the variation induced by such changes might be viewed as plausibly exogenous with respect to college choices, although of course this begs the question of why the political environment changed.[10]

As was shown in Figure 2.2 above, differences in support for higher education across States are highly persistent in recent years. Goldin and Katz (1999) show that these differences are in fact persistent over a much longer period of time, and they explain why:

> "To sum up, newer states with a high share of well-to-do families and scant presence of private universities in 1900 became the leaders in public higher education by 1930. They remain so today."

As Bound et al. (2004) point out, some of these differences across States might be related to other unmeasured differences in factors affecting college choices. For example, heavy industries requiring a lot of engineers and scientists might be located in places where conditions are favorable in terms of availability of natural resources, but unfavorable in that they happen to be populated by people who are skeptical about the value of higher education. In that case, the business community might push for more investment in public universities, and this would lead to a downward bias in estimates of the response to policy variables. On the other hand, Goldin and Katz (1999) argue that wealthier families are more likely to expect that their children will go to college, and indeed when they use automobiles per capita as a proxy for the level of wealth in the State, they find a positive relationship between wealth and public expenditures on higher education; this would lead to an upward bias in estimates

[8] Kane (2006) gives the example of California spending a lot on community colleges while also having low tuition.

[9] Card and Lemieux (2001) analyzed changes in college enrollment over the period 1968-1996, using a model of college participation that included tuition levels as one of the explanatory variables. The model includes State fixed effects, and also year fixed effects, so the effect of tuition is identified by differential changes in tuition over time within States – i.e. some States increased their tuition levels more or less quickly than others. The estimated effect of tuition is significant, but considerably smaller than the results in the previous literature (which used cross-section data, so that the effect is identified from differences in tuition levels across States at a point in time).

[10] Aghion et al. (2009) used a set of political instruments to distinguish between arguably exogenous variation in State expenditures on higher education and variation due to differences in wealth or growth rates across States. The model allows for migration, and it considers both innovation and imitation. Higher education investments affect growth in different ways depending on how close a State is to the "technology frontier". Each State is assigned an index measuring distance to the frontier, based on patent data. In States close to the frontier, the estimated effect of spending on research universities is positive, but the estimated effect is negative for States that are far from the frontier. The model explains this in terms of a tradeoff between using labor to innovate or to imitate.

of the response to policy variables. But although bias in one direction or the other cannot be ruled out, it seems reasonable to expect that differences in State policies arising from circumstances that prevailed many years ago would not be strongly related to unmeasured differences in determinants of college choices for recent cohorts (such as the NLSY79 cohort analyzed in this paper).

Bound et al. (2004) and Groen (2004) sidestep the issue of what causes changes in the number of college graduates in a State, and focus instead on the relationship between the flow of new graduates in a State and the stock of graduates working in that State some time later. They conclude that this relationship is weak, indicating that the scope for State policies designed to affect the educational composition of the labor force is limited. The empirical results presented here indicate that it is important to understand the sources of the variation in the flow of new graduates before drawing policy conclusions. This is discussed in Section 6 below.

4 A Life-Cycle Model of Expected Income Maximization

The empirical results in Kennan and Walker (2011) indicate that high school graduates migrate across States in response to differences in expected income. This paper analyzes the college choice and migration decisions of high-school graduates, using an extension of the dynamic programming model developed in Kennan and Walker (2011). The aim is to quantify the relationship between college choice and migration decisions, on the one hand, and geographical differences in college costs and expected incomes on the other. The model can be used to analyze the extent to which the distribution of human capital across States is influenced by State subsidies for higher education. The basic idea is that people tend to buy their human capital where it is cheap, and move it to where wages are high, but this tendency is substantially affected by moving costs.

Suppose there are J locations, and individual i's income y_{ij} in location j is a random variable with a known distribution. Migration and college enrollment decisions are made so as to maximize the present value of expected lifetime income. Let x be the state vector (which includes the stock of human capital, ability, wage and preference information, current location and age, as discussed below), and let a be the action vector (the location and college enrollment choices). In the basic dynamic discrete choice model[11], the utility flow is specified as $u(x,a) + \zeta_a$, where ζ_a is a random variable that is assumed to be iid across actions and across periods and independent of the state vector. It is assumed that ζ_a is drawn from the Type I extreme value distribution. Let $p(x' \mid x, a)$ be the transition probability from state x to x', if action a is chosen. The decision problem can be written in recursive form as

$$V(x,\zeta) = \max_a \left(v(x,a) + \zeta_a \right)$$

where

$$v(x,a) = u(x,a) + \beta \sum_{x'} p(x' \mid x, a) \bar{v}(x')$$

[11] See Rust (1994).

and

$$\bar{v}(x) = E_\zeta V(x, \zeta)$$

and where β is the discount factor, and E_ζ denotes the expectation with respect to the distribution of the vector ζ with components ζ_a. Then, using arguments due to McFadden (1974) and Rust (1994), we have

$$\exp(\bar{v}(x)) = \exp(\bar{\gamma}) \sum_{k=1}^{N_a} \exp(v(x,k))$$

where N_a is the number of available actions, and $\bar{\gamma}$ is the Euler constant. Let $\rho(x,a)$ be the probability of choosing a, when the state is x. Then

$$\rho(x,a) = \exp(v(x,a) + \bar{\gamma} - \bar{v}(x))$$

The function v is computed by value function iteration, assuming a finite horizon, T. Age is included as a state variable, with $v \equiv 0$ at age $T+1$, so that successive iterations yield the value functions for a person who is getting younger.

4.1 Nested and Sequential Choices

In the basic model, payoff shocks affecting enrollment and migration decisions are drawn independently from the Type I extreme value distribution. This is too restrictive: for example, enrollment choices might be more predictable than migration choices (or *vice versa*).

Suppose the choices are arranged in an array with m rows (for locations), and n columns (for enrollment choices). The model associates continuation values v_{ij} with row i and column j, and there are payoff shocks ζ_i associated with each row, and $\kappa\zeta'_j$ associated with each column, where the shocks are drawn independently from the Type I extreme value distribution, with $\kappa > 0$. Then if row i has been chosen, the column choice is determined by

$$j = \arg\max_k \left(v_{ik} + \zeta_i + \kappa\zeta'_k\right)$$

The (conditional) probability of choosing column j is

$$\rho(j \mid i) = \frac{\exp\left(\frac{v_{ij}}{\kappa}\right)}{\sum_{k=1}^{n} \exp\left(\frac{v_{ik}}{\kappa}\right)}$$

and the expected value of the row, $\bar{v}_i$, is given by

$$\exp(\bar{v}_i) = \exp(\kappa\bar{\gamma}) \left(\sum_{k=1}^{n} \exp\left(\frac{v_{ik}}{\kappa}\right)\right)^{\kappa}$$

If row i is chosen before the column shocks are realized (with the understanding that these shocks

will be realized before the column is chosen) then the row choice is determined by

$$i = \arg\max_s \left(\bar{v}_s + \zeta_s\right)$$

The probability of choosing row i is

$$\rho_0\left(i\right) = \frac{\exp\left(\bar{v}_i\right)}{\sum_{s=1}^{m} \exp\left(\bar{v}_s\right)}$$

and the expected value of the whole array is

$$\begin{aligned}
\bar{v}_0 &= \bar{\gamma} + \log\left(\sum_{s=1}^{m} \exp\left(\bar{v}_s\right)\right) \\
&= \bar{\gamma} + \log\left(\sum_{s=1}^{m} \exp\left(\kappa\bar{\gamma}\right)\left(\sum_{k=1}^{n} \exp\left(\frac{v_{sk}}{\kappa}\right)\right)^{\kappa}\right) \\
&= \left(1+\kappa\right)\bar{\gamma} + \log m + \log\left(\sum_{i=1}^{m}\left(\sum_{j=1}^{n} \exp\left(\frac{v_{ij}}{\kappa}\right)\right)^{\kappa}\right)
\end{aligned}$$

The choice probabilities are then given by

$$\begin{aligned}
\text{Prob}\left(d_{ij}=1\right) &= \rho\left(j \mid i\right)\rho_0\left(i\right) \\
&= \frac{\exp\left(\frac{v_{ij}}{\kappa}\right)}{\sum_{k=1}^{n} \exp\left(\frac{v_{ik}}{\kappa}\right)} \frac{\left(\sum_{k=1}^{n} \exp\left(\frac{v_{ik}}{\kappa}\right)\right)^{\kappa}}{\sum_{s=1}^{m}\left(\sum_{k=1}^{n} \exp\left(\frac{v_{sk}}{\kappa}\right)\right)^{\kappa}}
\end{aligned}$$

If $\kappa = 1$ (or if $m = 1$ or $n = 1$), this reduces to the standard logit formula for the choice probabilities.

The Nested Logit Model discussed by McFadden (1978) gives these same choice probabilities, but with a different interpretation: the continuation value associated with each choice is specified as $v_{ij} + y_{ij}$, where y_{ij} is a generalized extreme value random vector, with joint distribution function $F\left(y\right)$ given by

$$\begin{aligned}
F\left(y\right) &= \exp\left(-\sum_{i=1}^{m} Y_i\right) \\
Y_i^{\frac{1}{\kappa}} &= \sum_{j=1}^{n} \exp\left(-\frac{y_{ij}}{\kappa}\right)
\end{aligned}$$

subject to the restriction $0 < \kappa \leq 1$ (which ensures that the density function is non-negative).[12] In this interpretation all of the shocks are realized before any choices are made. In the present context, the period length is taken to be a year, and the timing of the location and enrollment

[12] See Börsch-Supan (1990). Note that the sequential choice interpretation allows $\kappa > 1$.

choices within the year is necessarily fuzzy, so various interpretations are possible, and each is just a rough approximation of the way that decisions are actually made. The estimated version of the model assumes that location choices are made before enrollment choices (but the reverse ordering gives similar results).

4.2 State Variables and Flow Payoffs

Let $\ell = \left(\ell^0, \ell^1\right)$ denote the current and previous location, let ω be a vector recording wage information at these locations, and let ξ denote current enrollment status (with the convention that $\xi = 0$ means that the individual is not enrolled in college, and otherwise ξ represents the college type). The state vector x consists of ℓ, ω, education level achieved so far, ability, parental education, family income, home location and age.[13]

The deterministic part of the flow payoff is specified as

$$u_h(x,j) = \alpha_0(e) + \alpha_1\left(1 - \mathbb{T}\left(\ell^0\right)\right) w\left(g, e, b, \ell^0, \omega, \xi\right) + \alpha_2 Y\left(\ell^0\right) + \alpha^H \chi\left(\ell^0 = h\right) - C_h\left(\ell^0, \xi\right) - \Delta(x,j)$$

Here the first term refers to consumption values associated with different education levels. The second term refers to after-tax wage income in the current location; the gross wage depends on age, schooling and ability, as discussed above, and $\mathbb{T}\left(\ell^0\right)$ is the tax rate in the current location. This is augmented by the amenity variable $Y\left(\ell^0\right)$. The parameter α^H is a premium that allows each individual to have a preference for their home location (χ denotes an indicator). The cost of attending a college of type ξ in location ℓ for a person whose home location is h is denoted by $C_h(\ell, \xi)$, with $C_h(\ell, 0) = 0$. The cost of moving from ℓ^0 to ℓ^j is represented by $\Delta(x,j)$.

4.3 Enrollment Decisions

In simple models of higher education choices, high school graduates choose whether to give up four or five years worth of earnings at high school wages in order to earn a college wage premium for the remaining forty years or so. In practice, the choices are more complicated. While many students enroll in college immediately after finishing high school, and stay in college continuously until they graduate, many others enroll in college after first spending some time in the labor force, or leave college without finishing a degree, either permanently or temporarily, on enroll in two-year colleges, with the possibility of subsequently transferring to a four-year college.[14] Accordingly, the model analyzed here treats college choices as the outcome of a sequence of decisions on whether to enroll in one of several types of college, with uncertainty about whether enrollment will lead to graduation with a degree.

The specification of the model involves the usual tradeoff between realism and computation; in particular, since there are many locations, and location is an essential state variable, it is necessary

[13] As in Kennan and Walker (2011), a limited (location) history approximation is used to reduce the size of the state space in a way that takes advantage of the low migration rates seen in the data.

[14] Agan (2014) presents a detailed description of the various paths taken by college students, using NLSY79 data.

to use a coarse specification of the other state variables so that the state space does not become too big. For this reason, there are just three levels of schooling: high school (12 or 13 years of schooling completed), some college (14 or 15 years) and college graduate (16 years or more).

In each period, there is a choice of whether to enroll in college. There are four types of college: community colleges, other public colleges and universities, and private colleges at two quality levels.[15] The college types differ with respect to tuition, State subsidies, financial aid, graduation probabilities, and psychic costs and benefits. Enrollment choices are influenced by ability, parental schooling and family income, represented by permanent state variables, which are restricted to just two values, high or low.[16]

4.4 Wages

The wage of individual i in location j at age g in year t is specified as

$$w_{ijt} = \mu_j(e_i) + \upsilon_{ij}(e_i) + G(e_i, X_i, g_i) + \varepsilon_{ijt}(e) + \eta_i$$

where e is schooling level, μ_j is the mean wage in location j (for each level of schooling), υ is a permanent location match effect, $G(e, X, g)$ represents the effects of observed individual characteristics, η is an individual effect that affects wages in the same way in all locations, and ε is a transient effect. The random variables η, υ and ε are assumed to be independently and identically distributed across individuals and locations, with mean zero. It is also assumed that the realizations of υ and η are seen by the individual (although $\upsilon_{ij}(e_i)$ is seen by i only after moving to j with education e_i).

The function G is specified as a piecewise-quadratic function of age g, with an interaction between ability and education:

$$G(e, b, g) = \begin{cases} \theta_e b + y_e^* - c_e (g - g_e^*)^2 & g \leq g_e^* \\ \theta_e b + y_e^* & g \geq g_e^* \end{cases}$$

where b is measured ability, y_e^* is the peak wage for education level e, and g_e^* is age at the peak. Thus both the shape of the age-earnings profile and the ability premium are specified separately for each level of education, with four parameters to be estimated ($\theta_e, y_e^*, c_e,$ and g_e^*).[17]

The relationship between wages and actions is governed by the difference between the quality of the match in the current location, measured by $\mu_j(e) + \upsilon_{ij}(e) + G(e, b, g)$, and the prospect of obtaining a better match in another location or at a higher level of schooling. The other components of wages have no bearing on migration or enrollment decisions, since they are added to the wage in the same way no matter what decisions are made.

[15] Fu (2014) uses a similar set of college types in her analysis of equilibrium in the college admissions market.

[16] Again, binary state variables are used here in order to keep the state space manageable.

[17] This is a version of the "flat-spot" specification developed by Heckman et al. (1998). In the present context, the model uses expected earnings at ages outside the range of the data to explain enrollment and migration choices at younger ages, and assuming that earnings are constant after the peak limits the possibility that the model might generate spurious explanations of these choices driven by implausible earnings at later ages.

4.4.1 Stochastic Wage Components

Since the realized value of the location match component υ is a state variable, it is convenient to specify this component as a random variable with a discrete distribution, and compute continuation values at the support points of this distribution. For given support points, the best discrete approximation $\hat{F}$ for any distribution F assigns probabilities so as to equate $\hat{F}$ with the average value of F over each interval where $\hat{F}$ is constant. If the support points are variable, they are chosen so that $\hat{F}$ assigns equal probability to each point.[18] Thus if the distribution of the location match component υ were known, the wage prospects associated with a move to State k could be represented by a distribution with equally weighted support points $\hat{\mu}_k + \hat{\upsilon}(q_r), 1 \leq r \leq R$, where $\hat{\upsilon}(q_r)$ is the q_r quantile of the distribution of υ, with

$$q_r = \frac{2r-1}{2R}$$

for $1 \leq r \leq R$. The distribution of υ is in fact not known, but it is assumed to be symmetric around zero. Thus for example with $R = 3$, the distribution of $\mu_j + \upsilon_{ij}$ in each State for each education level is approximated by a distribution that puts mass $\frac{1}{3}$ on μ_j (the median of the distribution of $\mu_j + \upsilon_{ij}$), with mass $\frac{1}{3}$ on $\mu_j \pm \upsilon^0$, where υ^0 is a parameter to be estimated.

Measured earnings in the NLSY are highly variable, even after controlling for education and ability. Moreover, while some people have earnings histories that are well approximated by a concave age-earnings profile, others have earnings histories that are quite irregular. In other words, the variability of earnings over time is itself quite variable across individuals. It is important to use a wage components model that is flexible enough to fit these data, in order to obtain reasonable inferences about the relationship between measured earnings and the realized values of the location match component. The wage components at each education level are specified as in Kennan and Walker (2011). The fixed effect η is assumed to be uniformly and symmetrically distributed around zero, with three points of support, so that there is one parameter to be estimated. The transient component ε should be drawn from a continuous distribution that is flexible enough to account for the observed variability of earnings. It is assumed that ε is drawn from a normal distribution with zero mean for each person, but with a variance that differs across people. Specifically, person i initially draws $\sigma_\varepsilon(i)$ from a uniform discrete distribution with two support points (which are parameters to be estimated), and subsequently draws ε_{it} from a normal distribution with mean zero and standard deviation $\sigma_\varepsilon(i)$, with ε_{it} drawn independently in each period.

4.5 College Costs

Aside from consumption values and expected income, all of the variables in the model that affect college choices do so by changing the costs associated with being in college. Earnings while enrolled in college are ignored. The college cost depends on ability, b, and on age, g (relative to an initial age g_0 which is set to 19). The cost also depends on resident and nonresident tuition rates, $\tau_r(\ell, \xi)$ and $\tau_n(\ell, \xi)$, expenditure on higher education, $y(\ell, \xi)$, financial aid (scholarships), $s(\ell, \xi)$, and parents'

[18] See Kennan (2006)

education and family income. Let d_m and d_f be indicators of whether the mother and the father have some college education, and let y_f be an indicator of whether family income is high or low. Let Ξ be the set of upper-tier colleges. The cost of attending a college of type ξ is specified as

$$\begin{aligned} C(\ell,\xi) &= \delta_0(\xi) + \delta_1(\xi)\tau(\ell,\xi) + \delta_2(\xi)E(\ell,\xi) + \delta_3 b + \delta_4 b\chi(\xi\in\Xi) + \delta_5 d_m + \delta_6 d_f + \delta_7 y + \delta_8(g-g_0) \\ &\quad + (\delta_9(\xi) + \delta_{10}b + \delta_{11}d_m + \delta_{12}d_f + \delta_{13}y_f)s(\ell,\xi) + (\delta_{14}+\delta_{15}b)\chi(e=1) + \delta_{16}\chi(\xi=\xi_{-1})\chi(e<2) \end{aligned}$$

where tuition is given by

$$\tau(\ell,\xi) = \chi(\ell=h)\tau_r(\ell,\xi) + \chi(\ell\neq h)\tau_n(\ell,\xi)$$

(with $\tau_r=\tau_n$ for private colleges). For each college type ξ, $\delta_0(\xi)$ measures the disutility of the effort involved in taking college courses (offset by the utility of life as a student); effort cost depends on ability (δ_3), especially in upper-tier colleges (δ_4), and the cost may be higher as students advance through college, especially for low-ability students ($\delta_{14}+\delta_{15}b$).[19] The tuition measures are averages over each college type within a State; it is assumed that the actual net tuition is a linear function of the State average tuition measures, and $\delta_1(\xi)$ represents the slope of this function, for each college type. Similarly, $\delta_2(\xi)$ measures the extent to which higher education expenditures reduce the cost of college, without specifying any particular channel through which this effect operates. The effect of scholarships is also measured separately for each college type, and in addition it depends on ability, parental education, and family income. The point here is that scholarships are largely allocated on the basis of merit or need; a college that has a large scholarship budget is more attractive (given tuition and expenditure levels), but the size of the scholarship budget is obviously more relevant for students who are more likely to be eligible for scholarships. Finally, persistence in enrollment choices is captured by the parameter δ_{16}.

4.6 Moving Costs

Moving costs are specified as in Kennan and Walker (2011). Let $D(\ell^0,j)$ be the distance from the current location to location j, and let $\mathbb{A}(\ell^0)$ be the set of locations adjacent to ℓ^0 (where States are adjacent if they share a border). The moving cost is specified as

$$\Delta(x,j) = \left(\gamma_0(e) + \gamma_1 D(\ell^0,j) - \gamma_2\chi\left(j\in\mathbb{A}(\ell^0)\right) - \gamma_3\chi\left(j=\ell^1\right) + \gamma_4 g - \gamma_5 n_j\right)\chi\left(j\neq\ell^0\right)$$

[19] In general it is not possible to distinguish between the nonpecuniary costs of college (δ_0) and the nonpecuniary benefits of having a college education (α_0). The income coefficient is identified by the migration component of the model. So the proportion who would choose college is known if there is no college cost, and if there is no difference between education levels except that college graduates earn more. Suppose the prediction is that the proportion going to college is 80%, and suppose that only 30% choose college in the data. The model might explain this by saying that going to college is costly. Alternatively, it might be explained by saying that there are nonpecuniary payoffs associated with the different education levels. The specification of costs and returns used here imposes an exclusion restriction that distinguishes one from the other: the transition probabilities are more favorable for high-ability people, but the nonpecuniary benefits of having a college education are the same for both types. This assumption is arbitrary. But the main point of the model is not to make these distinctions, but rather to estimate the responses to changes in the policy variables.

Thus the moving cost varies with education. The observed migration rate is much higher for college graduates than for high school graduates, and the model can account for this either through differences in potential income gains or differences in the cost of moving. The moving cost is an affine function of distance (which is measured as the great circle distance between population centroids). Moves to an adjacent location may be less costly (because it is possible to change States while remaining in the same general area). A move to a previous location may also be less costly, relative to moving to a new location. In addition, the cost of moving is allowed to depend on age, g. Finally, it may be cheaper to move to a large location, as measured by population size n_j.

4.7 Transition Probabilities

The state vector can be written as $x = (\tilde{x}, g)$, where $\tilde{x} = \left(e, \ell^0, \ell^1, x_v^0\right)$ and where x_v^0 indexes the realization of the location match component of wages in the current location. Let $q\left(e, \xi\right)$ denote the probability of advancing from education level e to $e+1$, for someone who is enrolled in a college of type ξ, with $q\left(e, 0\right) = 0$, and let $a = (j, \xi)$. The transition probabilities are as follows

$$p\left(x' \mid x\right) = \begin{cases} q\left(e,\xi\right) & \text{if} \quad j = \ell^0, \quad \tilde{x}' = \left(e+1, \ell^0, \ell^1, x_v^0\right), \quad g' = g+1 \\ 1 - q\left(e,\xi\right) & \text{if} \quad j = \ell^0, \quad \tilde{x}' = \left(e, \ell^0, \ell^1, x_v^0\right), \quad g' = g+1 \\ q\left(e,\xi\right) & \text{if} \quad j = \ell^1, \quad \tilde{x}' = \left(e+1, \ell^1, \ell^0, s_v\right), \quad g' = g+1, \quad 1 \leq s_v \leq n_v \\ 1 - q\left(e,\xi\right) & \text{if} \quad j = \ell^1, \quad \tilde{x}' = \left(e, \ell^1, \ell^0, s_v\right), \quad g' = g+1, \quad 1 \leq s_v \leq n_v \\ \frac{q(e,\xi)}{n} & \text{if} \quad j \notin \left\{\ell^0, \ell^1\right\}, \quad \tilde{x}' = (e+1, j, \ell^0, s_v), \quad g' = g+1, \quad 1 \leq s_v \leq n_v \\ \frac{1-q(e,\xi)}{n} & \text{if} \quad j \notin \left\{\ell^0, \ell^1\right\}, \quad \tilde{x}' = (e, j, \ell^0, s_v), \quad g' = g+1, \quad 1 \leq s_v \leq n_v \\ 0 & \text{otherwise} \end{cases}$$

5 Empirical Results

5.1 Data

The primary data source is the National Longitudinal Survey of Youth 1979 Cohort (NLSY79); data from the Census of Population are used to estimate State mean wages and parental income and education distributions, and data from the Integrated Postsecondary Education Data System (IPEDS) are used to measure tuition and college expenditures and financial aid. The NLSY79 conducted annual interviews from 1979 through 1994, and changed to a biennial schedule in 1994. The location of each respondent is recorded at the date of each interview, and migration is measured by the change in location from one interview to the next. Only the migration information from 1979 through 1994 is used here, but wage information is available (biennially) through 2013, and this is used in order to obtain better estimates of the lifetime wage profile.

In order to obtain a relatively homogeneous sample, only white non-Hispanic male high school graduates (or GED recipients) are included; the analysis begins at age 19. The (unbalanced) sample includes 12,895 annual observations on 1,281 men. Summary statistics on college enrollment for this sample are shown in Table 1.

Table 1: College Enrollment, NLSY

Enrollment Counts		
Public low	469	17%
Public high	1,497	53%
Private low	138	5%
Private high	737	26%
Subtotal	2,841	
Average years enrolled	3.7	
Not enrolled	10,054	
Total (person-years)	12,895	
Ever enrolled in college		
No	523	41%
In-State only	565	44%
Out-of-State only	98	8%
Both	95	7%
Total (persons)	1,281	

Wages are measured as total wage and salary income, plus farm and business income, adjusted for cost of living differences across States (using the ACCRA Cost of Living Index). The State effects $\{\mu_j(e)\}$ are obtained from 1980 and 1990 Census data, using median wage regressions with year and age and State dummies, applied to white males who have recently entered the labor force (so as to minimize selection effects due to migration).[20] The tax rate $\mathbb{T}\left(\ell^0\right)$ is measured as the sum of federal and State average tax rates for 1984, as calculated by the NBER TAXSIM model.[21]

5.1.1 Tuition and Subsidies

In the model, each State has one representative college of each type[22], and all of these colleges are available choices for everyone.[23] Tuition rates were estimated by computing enrollment-weighted averages of "sticker prices" for each college type, using IPEDS data for 1984. Students attending college in their home State are assumed to pay tuition at the resident rate, while others pay the non-resident rate, allowing for a few reciprocity agreements across States.[24] The home State is defined as the State in which the individual last went to high school.

State subsidies to higher education might affect either the cost or the quality of education.

[20]Since wages include location match effects, if mean State wages were estimated using data for everyone currently working in the State, the estimated match effect distribution would be biased.

[21]See http://users.nber.org/~taxsim/allyup/.

[22]There are a few exceptions: there are no private colleges in Wyoming (aside from Wyoming Technical Institute, a for-profit operation of dubious repute), and there are no upper-tier private colleges in Montana, Nevada and South Dakota. Thus these alternatives are excluded from the choice set in the dynamic programming model.

[23]This does not mean that every high school graduate is free to choose Harvard. There are 43 colleges in Massachusetts that are classified as upper-tier (including Harvard), and the assumption is that every high school graduate can get into at least one of these colleges.

[24]Minnesota has tuition reciprocity agreements with Wisconsin and with North and South Dakota; there is a similar agreement between Oregon and Washington State.

For example, given the level of tuition, the cost of attending college is lower if there is a college within commuting distance, and the cost of finishing college is higher if graduation is delayed due to bottlenecks in required courses. From the point of view of an individual student, an increase in tuition paid by other students has much the same effect as an increase in subsidies, in the sense that it increases the resources available for instruction and student support services. But because tuition also acts as a price, it seems more informative to model the effect of direct subsidies, holding tuition constant. This means that the effect of tuition should not be interpreted as a movement along a demand curve, since a college that charges high tuition, holding subsidies constant, can use the additional tuition revenue to improve the quality of the product, or to reduce other components of college costs.

Subsidy measures were constructed by adding federal, State and local appropriations and grants over all public colleges in the 1984 IPEDS file, by State, and by college level, the lower level being defined as community colleges, and the upper level as all other public colleges. Similarly, the financial aid variables measure total expenditures on scholarships, by State and college level.[25] Since these expenditure aggregates involve populations of different sizes, the expenditure and financial aid figures are divided by the number of potential students, measured as the number of high school graduates in the State aged 22-36 in the 1990 Census. Summary statistics are shown in Table 2

5.1.2 College Choices

As is well known, there is a strong relationship between college choices and parental education levels. For the sample used here, this relationship is summarized in Table 3, for low-ability and high-ability students, where the ability measure is an indicator of whether the AFQT percentile score is above or below the median in the full sample (which is 63).

For example, if both parents went to college, there is a 52% chance that their sons will graduate from college, and this rises to 64% if the son is in the top half of the distribution of AFQT scores. There is also a strong relationship between AFQT scores and college choices, but note that sons whose parents went to college are much more likely to have high AFQT scores.

Transition rates for the NLSY sample are shown in Table 4. These are treated as transition probabilities, and held fixed when the the model is estimated.

5.2 College Choices and Migration

Table 5 gives the main empirical results. The parameters of the wage process are estimated separately, using the most recent data (including the biennial interviews)[26]; these parameters, which are shown in the right panel of Table 5, are treated as known when estimating the other parameters

[25] These data were obtained from the IPEDS finance file for 1984 (nces.ed.gov/ipeds/datacenter/data/F1984_Data_Stata.zip); the expenditure variable includes expenditures on Instruction, Research, Public service, Academic support (excluding libraries), Student services, Institution support, and Educational Mandatory Transfers. The financial aid variable includes Scholarships (unrestricted) and Scholarships (restricted).

[26] The wage unit is $\$10,000$ (at 1983 prices).

Table 2: **Wage Differentials and College Costs**

	Mean	S.D.	Min	Max
Earnings ($1983)				
High School (age 20)	7,856	871	5,824	10,196
Some College (age 22)	9,966	982	7,451	11,809
College Graduate (age 24)	13,984	1,271	9,345	16,174
Tuition				
Public, low, Resident	663	280	86	1,422
Public, low, Nonresident	1,830	738	555	3,742
Public, high, Resident	1,224	398	471	2,553
Public, high, Nonresident	3,166	903	1,532	6,181
Private, Low	3,767	927	1,438	5,749
Private, High	5,197	1,765	1,518	9,166
Expenditure (per potential student)				
Public, low	111	92	13	402
Public, high	679	252	227	1,474
Private, Low	54	53	2	311
Private, High	218	224	2	898
Financial Aid (per potential student)				
Public, low	13.0	8.8	2.8	41.0
Public, high	51.6	24.8	16.3	149.3
Private, Low	13.1	10.9	0.7	59.2
Private, High	33.2	30.5	0.1	136.3

Table 3: **Ability, Parents' Education and Schooling**

Neither Parent went to College				
	High School	Some College	College	Total
Years	12-13	14-15	16+	
Low Ability	375	33	34	442
	84.8%	7.5%	7.7%	62.3%
High Ability	128	56	84	268
	47.8%	20.9%	31.3%	37.7%
Total	503	89	118	710
	70.8%	12.5%	16.6%	
Both Parents went to College				
Low Ability	41	19	19	79
	51.9%	24.1%	24.1%	29.7%
High Ability	24	44	119	187
	12.8%	23.5%	**63.6%**	70.3%
Total	65	63	138	266
	24.4%	23.7%	51.9%	

Table 4: **College Transition Rates**

			Low AFQT		**High AFQT**	
	Initial Grade		12-13	14-15	12-13	14-15
	e		0	1	0	1
	Next Grade		14-15	16	14-15	16
	e		1	2	1	2
		ξ				
Public	Lower-Tier	1	24.4%	13.2%	32.9%	6.3%
Public	Upper-Tier	2	45.1%	35.6%	56.7%	34.6%
Private	Lower-Tier	3	44.4%	18.2%	62.9%	41.7%
Private	Upper-Tier	4	41.3%	29.5%	57.5%	35.7%

governing college choice and migration decisions[27]. The estimates of the parameters governing migration decisions are similar to the estimates in Kennan and Walker (2011). The estimated income coefficient in this model reflects both migration and college choice decisions; as in the migration model, the effect is highly significant. Ability and parental education levels have strong effects on college costs (as would be expected, given the data in Table 3). The sequential structure of the payoff shocks substantially improves the model fit, and the estimate of κ indicates that migration decisions are much less predictable than enrollment decisions. The estimated moving costs are decreasing in the level of education, reflecting the positive relationship between education and migration rates in the data. The age coefficients for both moving and enrollment costs are quite significant. If these coefficients were zero, the model could still explain why younger people are more likely to enroll in college, just as they are more likely to move: these are both investment decisions, and if the net return is positive, it is better to invest sooner rather than later. The estimates indicate that this human capital explanation is insufficient to fully explain why observed enrollment and migration rates are decreasing in age.

[27]Surprisingly, the direct effect of the (binary) AFQT score is weak (conditional on the education level). Ability is of course strongly correlated with earnings, but the estimated earnings process attributes this almost entirely to a strong relationship between ability and educational attainment.

Table 5: **College Enrollment Choices and Migration, White Males**

Utility Parameters			
		$\hat{\theta}$	$\hat{\sigma}_\theta$
A. Consumption value of education			
Some College	$\alpha_0(1)$	-0.256	0.053
College Graduate	$\alpha_0(2)$	-0.010	0.171
B. Flow Utility			
Income	α_1	0.098	0.010
Climate	α_2	0.010	0.004
Home Premium	α^H	0.164	0.009
C. Moving Cost			
High School	$\gamma_0(1)$	4.613	0.252
Some College	$\gamma_0(2)$	4.090	0.268
College Graduate	$\gamma_0(3)$	4.122	0.283
Distance	γ_1	0.312	0.062
Adjacent Location	γ_2	0.886	0.080
Previous Location	γ_3	0.164	0.009
Age effect	γ_4	2.305	0.113
Population	γ_5	0.083	0.010
D. College Cost			
Disutility, college: Pub lo	$\delta_0(1)$	1.113	0.273
Disutility, college: Pub hi	$\delta_0(2)$	-0.084	0.376
Disutility, college: Pvt lo	$\delta_0(3)$	1.800	0.445
Disutility, college: Pvt hi	$\delta_0(4)$	0.927	0.395
Tuition: Public lo	$\delta_1(1)$	1.138	0.373
Tuition: Public hi	$\delta_1(2)$	2.146	0.187
Tuition: Private lo	$\delta_1(3)$	-1.599	0.561
Tuition: Private hi	$\delta_1(4)$	-0.087	0.224
Spend/Student: Pub lo	$\delta_2(1)$	-17.703	2.121
Spend/Student: Pub hi	$\delta_2(2)$	0.656	0.763
Spend/Student: Pvt lo	$\delta_2(3)$	11.460	29.693
Spend/Student: Pvt hi	$\delta_2(4)$	-7.160	2.899
Ability effect on cost	δ_3	-0.279	0.160
Ability×upper tier	δ_4	-0.312	0.074
Mother's education	δ_5	0.069	0.032
Father's education	δ_6	-0.125	0.033
Family Income	δ_7	-0.067	0.029
Age effect	δ_8	0.279	0.160
Financial Aid: Pub lo	$\delta_9(1)$	-17.762	37.030
Financial Aid: Pub hi	$\delta_9(2)$	21.243	9.867
Financial Aid: Pvt lo	$\delta_9(3)$	-119.044	128.454
Financial Aid: Pvt hi	$\delta_9(4)$	32.017	23.137
Ability×aid	δ_{10}	-41.933	7.302
Mother-ed×aid	δ_{11}	-41.628	6.953
Father-ed×aid	δ_{12}	-3.550	7.191
Family income×aid	δ_{13}	-7.318	7.100
extra cost, upper ed level	δ_{14}	0.868	0.257
hi ability, upper ed	δ_{15}	0.483	0.174
enrollment persistence	δ_{16}	-1.421	0.102
E. Payoff shocks			
Enroll/migrate shocks	κ	0.504	0.032

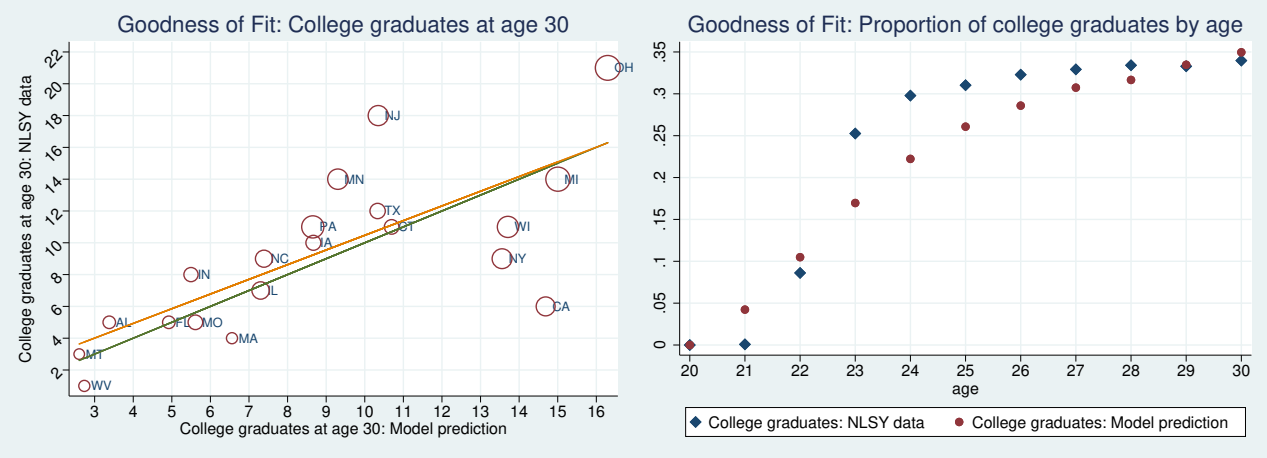
	Wage Parameters					
	High School		Some College		College	
	$\hat{\theta}$	$\hat{\sigma}_\theta$	$\hat{\theta}$	$\hat{\sigma}_\theta$	$\hat{\theta}$	$\hat{\sigma}_\theta$
Peak Wage	1.801	0.015	2.400	0.041	2.367	0.041
Age at Peak	38.334	0.376	44.961	1.327	48.922	0.713
Curvature	1.175	0.053	0.943	0.120	1.285	0.073
AFQT	0.139	0.024	0.124	0.044	0.150	0.040
Location match	0.424	0.009	0.911	0.022	0.892	0.017
Transient s.d. 1	0.593	0.002	0.791	0.007	0.789	0.005
Transient s.d. 2	2.651	0.009	3.831	0.053	4.484	0.030
Individual Effect	1.138				0.012	

For public colleges, higher tuition has a strong negative effect on enrollment, and expenditure per (potential) student has a strong positive effect for community colleges (but the expenditure effect is insignificant for other public colleges). There is considerable variation in tuition levels for private colleges, but since this variation is not determined at the State level, the effect of differences in private college tuition cannot easily be inferred from location choices, as is done here for public colleges.

5.3 Goodness of Fit

The ability of the model to fit the data might be measured in various ways. Two such measures are shown in Figure 5.1. The figure on the left shows the actual and predicted numbers of college graduates at age 30 (for States with at least 10 observations at this age), with a regression line and a 45-degree line superimposed. Although the fit is far from perfect, the model gives a reasonably good description of the cross-State variation in graduate rates. The figure on the right shows the predicted and actual timing of college graduation. Again, the fit is reasonable, although the model misses the jump in graduation rates around age 23, as might be expected given that the model does not allow for heterogeneous transition probabilities.[28]

Figure 5.1: Goodness of Fit

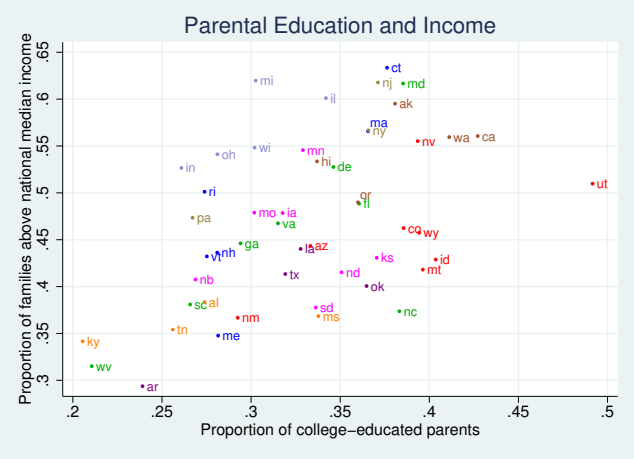


6 Effects of Changes in Tuition Levels and State Expenditures

The results in Table 5 indicate that college enrollment decisions are affected by tuition and expenditure and financial aid levels, while expected income differences affect both enrollment and migration decisions. The question then is whether changes in State policies regarding tuition and expenditures

[28] Suppose 50% of college students will eventually graduate. In the model, this is achieved by taking repeated draws from the transition probability distribution, until two successful transitions occur. In the data, there are some people who will graduate with much higher probability, and they will get through the system quickly, and there will be others who never get through. This is a kind of mover-stayer problem, and it could be handled by allowing for two unobserved types (in addition to the two observed ability types). But this would make the model much more complicated. In the model without this kind of heterogeneity, the transition probability is such that the right number of people get through in the end, but this is achieved by spreading the process over more periods.

Figure 6.1: Parental Education and Family Income



have long-term effects on the educational composition of the State's labor force, as opposed to transient effects that are undone by migration, as suggested by Bound et al. (2004). The main point of the model is that it can answer questions of this kind.

Suppose for example that Michigan reduces tuition, or increases expenditures. The effects of such changes are presumably small for high school graduates in Alaska or Louisiana, but perhaps not so small for students from Michigan, or neighboring States. Moreover, the effects depend on individual characteristics. The model has 800 types, classified by home location, and by four binary characteristics (ability, family income, and mother's and father's education). In order to estimate the effects of changes in college costs (or wages) it is necessary to use the respective value functions to compute the responses for each type, and then construct a suitably weighted average over types. The main complication here is that parental education and family income vary considerably across States. To deal with this, data from the 1970 Census were used to identify households with children aged 5-13 (corresponding to the ages of the individuals in the NLSY data), and the family income and parental education data for these children were then tabulated, by State. The results of these tabulations are illustrated in Figure 6.1.

The proportion of high-ability types in each State is estimated using the AFQT scores and the parental education and household income data in the NLSY sample. Surprisingly, the binary family income variable doesn't help explain ability differences (there is a slight relationship if parental education variables are excluded, but there is no effect given parental education). And if just one parent has been to college, it doesn't really matter which one. The estimated ability proportions are shown in Table 6. These proportions are applied to the parental education data from the 1970 Census, meaning that the proportion of high-ability types is higher for States on the right side of Figure 6.1.

Table 6: **Ability and Parental Education**

	Ability		
Parents went to college?	Low	High	Percentage
Neither	442	268	37.75
One	128	177	58.03
Both	79	187	70.30
Total	649	632	49.34

6.1 Equilibrium Wage Effects

If policy changes affect the educational composition of the labor force, relative wages might also be affected. It is true that in the long run, if each individual State is considered as a small open economy, the factor price equalization theorem implies that wages are invariant to local changes in the supply of labor at different education levels [29]. But there is an extensive literature that considers the supply and demand for labor without dealing with the general equilibrium effects underlying the factor price equalization theorem, analyzing the empirical relationship between the return to education and the relative supplies of workers at different education levels.[30] Building on this literature, Fortin (2006) estimated that the "own-cohort relative supply effect" is about -0.2, meaning that a 10% increase in the relative supply of college graduates in a State generates a 2% reduction in the college wage premium, within cohorts spanning 10 years. The policy effect estimates presented below are based on the assumption that wages do not change, as would be true if the assumptions of the factor price equalization theorem are empirically valid. But in case wages do actually change, it is clearly of interest to determine whether such changes are large enough to undo the direct effects of policy changes, and the results from Fortin (2006) will be used for this purpose.

6.2 Simulation Results: Effects of Changes in Policy Variables

The evolution of the population distribution in the model is computed by iterating the transition matrix of the Markov chain on the state space. The model specifies choice probabilities $\rho(x, a)$, where x is the state vector, and a is the choice variable; the next state x' is then determined by the transition probabilities $q(x, a, x')$. There is a frequency distribution $p(x)$ over current states, and the model implies a transition matrix T from $p(x)$ to $p'(x)$ given by

$$T(p)(x) = \sum_{t \in X} p(t) \sum_{a \in A} \rho(t, a)\, q(t, a, x)$$

The effects of changes in the policy variables are computed by first iterating the transition matrix implied by the values of the policy variables used in the estimation, and then doing the same thing

[29] Similarly, factor price equalization implies that immigration does not affect wages; see Kennan (2013)

[30] See Katz and Murphy (1992) and the many thousands of references thereto (scholar.google.com/citations?user=MAqEdFMAAAAJ).

for the new values of the policy variables, and comparing the population distributions.[31]

Table 7: **Effects of Policy Changes: Michigan**

		Population at Age 36						Revenue	
		Current Location			**Home Location**			Taxes	Tuition
		Graduates	Some College	High School	Graduates	Some College	High School		
Increase (20%)									
Tuition	Michigan	**-11.2%**	**-2.4%**	**6.2%**	**-12.9%**	**-2.4%**	**6.5%**	**25.2**	**2.1**
	elsewhere	-0.16%	0.00%	0.05%	-0.03%	0.01%	0.01%	-82.1	2.5
Spending	Michigan	**0.03%**	**0.42%**	**-0.16%**	**0.02%**	**0.45%**	**-0.17%**	**-10.3**	**2.6**
	elsewhere	0.001%	0.016%	-0.005%	0.001%	0.004%	-0.002%	2.3	1.2
Decrease (20%)									
Tuition	Michigan	**12.7%**	**1.7%**	**-6.6%**	**14.3%**	**1.6%**	**-6.9%**	**-13.6**	**-45.0**
	elsewhere	0.19%	-0.02%	-0.07%	0.05%	-0.02%	-0.02%	76.4	-4.6
Spending	Michigan	**0.11%**	**-1.36%**	**0.43%**	**0.12%**	**-1.47%**	**0.45%**	**9.2**	**-0.3**
	elsewhere	0.001%	-0.015%	0.004%	0.000%	-0.004%	0.001%	-1.1	-1.0

Some illustrative results are shown in Table 7, taking Michigan as an example; results for some other large States are shown in Tables 8 and 9. The population distributions over locations and educational attainment are compared at age 36, with location defined alternatively as current or home location. These tables also show the effects on policy changes on State revenues, computed as the present values of tax and tuition revenues, from age 19 to age 60, with a discount factor of .95 (as in the estimated model).

Table 7 shows that 20% changes in tuition or expenditures generate substantial (and roughly symmetric) changes in the proportions of college-educated men. The tuition effects are generally larger than the expenditure effects, although expenditure changes do seem to influence the number of people who have some college education, with generally smaller effects on the number of graduates (which is consistent with the coefficient estimates in Table 5). The estimated effects vary considerably across States. To some extent this variation arises because a 20% change in tuition or expenditures corresponds to different dollar amounts, depending on the initial level. But there is still considerable cross-State variation even if the changes are rescaled to represent equal dollar amounts in each State. The most striking result is that, contrary to the findings in Bound et al. (2004), changes in the policy variables have substantial long-term effects on the educational composition of the local population many years after the fact, despite some leakage due to migration.

Although the effects of tuition reductions are not exactly revenue-neutral, the loss of tuition revenue is largely offset by the increased tax revenue generated by a more educated labor force. Indeed the estimates in Table 8 indicate that although net revenue generally falls somewhat, there is actually an increase in net revenue in the case of Pennsylvania, while the effects are revenue-neutral for Michigan and Illinois.

[31] Note that there is no need to simulate actual choices, so there is no simulation error in these calculations (aside from rounding errors arising from repeated multiplication of large probability matrices that have some very small elements

Table 8: **Effects of Tuition Reductions**(20%)

	Population at Age 36						Revenue (present values)			
	Current Location			Home Location			Taxes		Tuition	
	Graduates	Some College	High School	Graduates	Some College	High School	Δ%	$	Δ%	$
California	3.6%	-1.0%	-3.6%	3.5%	-1.2%	-3.9%	0.19%	56.3	-12.3%	-149.2
Florida	5.9%	0.4%	-4.1%	6.9%	0.1%	-4.9%	0.27%	16.2	-10.0%	-44.8
Illinois	9.4%	1.0%	-6.1%	10.8%	0.9%	-6.5%	0.15%	26.7	-6.9%	-71.9
Massachusetts	3.2%	0.1%	-3.2%	3.4%	0.1%	-3.4%	-0.12%	-14.6	-8.8%	-45.6
Michigan	12.7%	1.7%	-6.6%	14.3%	1.6%	-6.9%	-0.08%	-13.6	-4.3%	-45.0
New York	7.6%	0.2%	-5.8%	8.1%	0.0%	-6.2%	0.05%	17.6	-8.7%	-165.4
Pennsylvania	13.4%	3.5%	-6.1%	17.9%	3.8%	-6.4%	-0.17%	-33.4	2.0%	19.9
Texas	4.2%	0.4%	-2.8%	4.4%	0.2%	-3.0%	0.33%	48.6	-13.3%	-80.5
Wisconsin	9.6%	1.1%	-5.1%	10.6%	1.0%	-5.4%	-0.13%	-11.3	-6.9%	-34.7

Table 9: **Effects of Increases in Expenditures**(20%)

	Population at Age 36						Revenue	
	Current Location			Home Location			present values	
	Graduates	Some College	High School	Graduates	Some College	High School	Taxes	Tuition
California	4.2%	7.5%	-10.3%	4.4%	7.9%	-11.4%	-153.4	17.9
Florida	5.5%	7.4%	-7.8%	7.3%	8.6%	-9.6%	-22.9	51.9
Illinois	9.41%	0.98%	-6.12%	10.79%	0.87%	-6.53%	26.7	-71.9
Massachusetts	0.03%	0.42%	-0.16%	0.02%	0.45%	-0.17%	-3.0	1.2
Michigan	0.03%	1.53%	-0.56%	0.05%	1.64%	-0.58%	-10.3	2.6
New York	0.21%	1.40%	-0.70%	0.25%	1.49%	-0.74%	-22.9	51.9
Pennsylvania	-0.27%	0.26%	0.03%	-0.36%	0.28%	0.03%	-1.3	-4.8
Texas	4.19%	0.37%	-2.82%	4.45%	0.23%	-3.04%	48.6	-80.5
Wisconsin	9.56%	1.06%	-5.10%	10.64%	0.96%	-5.43%	-1.3	-0.1

Bound et al. (2004) found a fairly weak relationship between flows of new college graduates and subsequent stocks of graduates in the labor force at the State level, suggesting that there is relatively little scope for State policies that aim to increase the proportion of college graduates in the State labor force by investing more in the State's public colleges. The interpretation of this finding is that increases in the flow of college graduates generated by tuition reductions or expenditure increases in the State do not have much effect on long-term stocks because workers are mobile (and college graduates are particularly mobile). One problem with this interpretation is that there is no analysis of what caused the flow increase, and there are good reasons to expect that the proportion of the flow increase that "sticks" in the State is not invariant with respect to alternative causes of the increase. One example is that if the increased flow of new graduates was generated by attracting students from other States, then it is likely that many of these students would return home or move elsewhere after graduation, whereas an increase in the number of students from this State would be associated with a strong tendency to remain in the State after graduation. In contrast, this paper analyzes specific policy changes, keeping track of the effects of these changes on the choices made by individuals who differ in various respects, and in particular allowing for migration decisions that are strongly affected by individuals' home locations. This gives sharper conclusions, especially with respect to the leakage of college graduates due to migration. Indeed, the results indicate that the leakage due to migration is negligible.

This analysis can be illustrated by considering the effects of changes in resident tuition rates, with non-resident rates held constant, and vice versa. Table 10 shows the results of simulating these effects. Clearly, the increases in the proportion of college graduates in a State's labor force shown in Table 6 are almost entirely attributable to the effects of tuition reductions for residents of that State. Reductions in nonresident tuition do lead to increases in the proportion of college graduates, but these effects are very small. One can also ask whether the additional graduates tend to stay in the State that reduced tuition rates. Again, the answer depends heavily on whether the change was directed at residents or nonresidents; in the case of a change in resident tuition, the proportion of stayers is high, while in the nonresident case, the proportion is much lower (especially for non-peripheral States).

6.2.1 National Tuition Rates

Given the sizable effects of tuition changes, and the cross-State dispersion in tuition levels, it is natural to consider the effects of a national tuition policy. For example, tuition levels in California are lower than in all other States except for Texas, and the proportion of college graduates in California is high. If all States set tuition at the California (resident) level, would there be a substantial increase in the proportion of college graduates in the country as a whole? And if so, how costly would this be, in terms of lost tuition revenue, net of the increase in tax revenue?

Answers to these questions are shown in Table 11. Tuition levels in public universities and community colleges in all States are set at the California in-State levels, and income tax rates are

associated with unlikely choices).

Table 10: **Effects of Resident and Nonresident Tuition Reductions**(20%)

	Population at Age 36				Revenue (present values)			
	Resident		Nonresident		Resident		Nonresident	
	Graduates	Stayers	Graduates	Stayers	Taxes	Tuition	Taxes	Tuition
California	3.23%	81.9%	0.42%	54.3%	-6.5	-169.5	62.7	20.3
Florida	4.89%	83.7%	1.02%	52.3%	-7.9	-51.2	24.0	6.4
Illinois	3.64%	77.8%	0.16%	42.2%	-8.5	-83.0	35.2	11.0
Massachusetts	1.39%	72.1%	0.05%	36.6%	-30.9	-50.0	16.3	4.4
Michigan	12.16%	73.8%	0.53%	35.1%	-39.5	-59.3	25.9	14.3
New York	2.67%	75.7%	0.11%	38.9%	-52.3	-170.3	69.9	4.9
Pennsylvania	5.72%	73.3%	0.16%	35.1%	-54.0	9.1	20.6	10.8
Texas	1.13%	84.1%	0.13%	59.3%	7.2	-78.9	41.4	-1.6
Wisconsin	3.39%	74.1%	0.18%	34.2%	-27.3	-44.4	16.0	9.7

Note: "Stayers" means the proportion of the national increase in graduates found in the State where the tuition rate changed.

Table 11: **Effects of Revenue-Neutral National Tuition Reductions**(20%)

	Population at Age 36							
	Current Location			Home Location				
		Some	High		Some	High		
	Graduates	College	School	Graduates	College	School	Taxes	Tuition
CA	10.8%	-4.8%	-6.8%	6.6%	-5.3%	-5.3%	766	-90
FL	28.6%	-1.4%	-15.1%	21.1%	-1.9%	-13.8%	235	-197
IL	41.0%	-4.4%	-22.4%	43.2%	-4.6%	-22.8%	312	-413
MA	12.1%	-6.0%	-13.5%	14.8%	-4.9%	-13.1%	-134	-226
MI	57.1%	-3.3%	-25.9%	62.6%	-3.3%	-26.5%	68	-449
NY	32.0%	-5.8%	-22.5%	32.6%	-5.8%	-22.8%	293	-898
PA	80.2%	3.1%	-30.1%	98.8%	3.8%	-30.8%	111	-374
TX	9.2%	-4.2%	-2.0%	3.2%	-3.9%	-0.4%	372	-6
WI	36.6%	-3.7%	-17.4%	38.5%	-3.7%	-17.5%	42	-175
US	31.2%	-3.2%	-17.2%	31.2%	-3.2%	-17.2%	6167	-6158

adjusted so that the reduction in tuition revenue is offset by an increase in tax revenue. The results are quite striking. When California tuition levels are applied uniformly in all States, the proportion of college graduates increases by a factor of about one-third. The increase is distributed unevenly, with much larger increases in high-tuition States (like Pennsylvania), and much smaller increases in low-tuition States (like Texas), but the overall result is a dramatic increase in the number of college graduates.[32] Moreover, the tax increase required to offset the lost tuition revenue is very small: a 1% surcharge yields a small surplus.[33]

6.2.2 General Equilibrium Effects

The estimates in Table 11 assume that changes in the proportion of college graduates have no effect on wages. As was mentioned in Section 6.1 above, the results from Fortin (2006) can be used to obtain a rough idea of the quantitative importance of this assumption. Suppose the ratio of college graduates to high school graduates increases by 60% (which is about the right magnitude according to Table 11, if the some college group is ignored). The Fortin estimate says that a 10% increase in the relative supply means a 2% reduction in the college wage premium, and if this is extrapolated, the implication is that the wage premium falls by about 12%. It turns out that if high school wages are increased by 6% and college wages are decreased by 6%, the results in Table 11 are not much affected – there is still a very large increase in the proportion of college graduates.

7 Conclusion

There are big differences across States in the extent to which higher education is subsidized, and the State subsidies are apparently motivated to a large extent by a perceived interest in having a highly educated labor force. There are also substantial wage differences across States, and previous work has found that these generate sizable supply responses in NLSY data. In the absence of moving costs, income maximization implies that human capital should be acquired in locations where it is cheap, and subsequently deployed in labor markets where the return is high, implying that differences in State policies have little effect on the educational composition of the labor force. But moving costs are important, and most people have a strong preference for their home location. The paper uses a dynamic programming model of income-maximizing college enrollment and migration decisions, allowing for locational preferences and moving costs, and uses it to estimate the extent to which the educational composition of the labor force is affected by inter-State differences in higher education financing policies.

The model is estimated on NLSY79 data for white male high school graduates. The results suggest that more generous subsidies actually do have significant effects on college enrollments,

[32] There is actually a substantial increase in the proportion of college graduates among California residents, even though there is no change in tuition levels there, because some California residents take advantage of the reduced tuition levels in other States.

[33] The average tax rate in the baseline simulation is 8.92%; the results in Table 11 were generated by multiplying the tax rate in each State by a factor of 1.01, meaning that the average tax rate rises to 9.01%

particularly when the subsidies take the form of tuition reductions. Moreover, the effects of higher education financing policies on the educational level of the local labor force are apparently long-lasting: the extent to which they are dissipated through migration is very slight. A particularly striking implication of the estimated model is that a national tuition policy, set at a uniformly low level (corresponding to the actual level for California residents), would yield a very large increase in the proportion of college graduates, and the resulting increase in income tax revenue would almost completely offset the loss in tuition revenue.

References

Agan, Amanda Y., "Disaggregating the Returns to College," January 2014. Unpublished, Princeton University.

Aghion, P., L. Boustan, C. Hoxby, and J. Vandenbussche, "The Causal Impact of Education on Economic Growth: Evidence from the United States," 2009. Unpublished, Harvard University.

Bishop, Kelly C., "A Dynamic Model of Location Choice and Hedonic Valuation," November 2012. Unpublished, Washington University in St. Louis.

Börsch-Supan, Axel, "On the compatibility of nested logit models with utility maximization," *Journal of Econometrics*, 1990, *43* (3), 373–388.

Bound, John, Jeffrey Groen, G.Gábor Kezdi, and Sarah Turner, "Trade in university training: cross-state variation in the production and stock of college-educated labor," *Journal of Econometrics*, 2004, *121* (1-2), 143 – 173. Higher education (Annals issue).

Card, David and Alan B. Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *The Journal of Political Economy*, 1992, *100* (1), 1–40.

_ **and Thomas Lemieux**, "Dropout and Enrollment Trends in the Postwar Period: What Went Wrong in the 1970s?," in Jonathan Gruber, ed., *An Economic Analysis of Risky Behavior Among Youth*, Chicago: University of Chicago Press, January 2001, pp. 439–482.

Dynarski, Susan, "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance," *National Tax Journal*, 2000, *53* (3 Part 2), 629–662.

Dynarski, Susan M., "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion," *The American Economic Review*, 2003, *93* (1), 279–288.

Fortin, Nicole M., "Higher-Education Policies and the College Wage Premium: Cross-State Evidence from the 1990s," *American Economic Review*, September 2006, *96* (4), 959–987.

Fu, Chao, "Equilibrium Tuition, Applications, and Admissions and Enrollment in the College Market," *Journal of Political Economy*, 2014, *122* (2), 225–281.

Gemici, A., "Family migration and labor market outcomes," January 2011. Unpublished, New York University.

Goldin, Claudia and Lawrence F. Katz, "The Shaping of Higher Education: The Formative Years in the United States, 1890 to 1940," *Journal of Economic Perspectives*, 1999, *13* (1), 37–62.

Groen, Jeffrey A., "The effect of college location on migration of college-educated labor," *Journal of Econometrics*, 2004, *121* (1-2), 125 – 142. Higher education (Annals issue).

Heckman, James J, Lance Lochner, and Christopher Taber, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1998, *1* (1), 1 – 58.

Kane, Thomas J., "Public Intervention in Post-Secondary Education," in E. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*, Vol. 2, Amsterdam, New York and Oxford: Elsevier Science, North-Holland, 2006, pp. 1369 – 1401.

_ , "Evaluating the Impact of the D.C. Tuition Assistance Grant Program," *Journal of Human Resources*, 2007, *XLII* (3), 555–582.

Katz, Lawrence F and Kevin M Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand factors," *Quarterly Journal of Economics*, 1992, *107* (February), 35–78.

Keane, Michael P. and Kenneth I. Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, June 1997, *105* (3), 473–522.

_ **and** _ , "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review*, 2001, *42* (4), 1051–1103.

Kennan, John, "A Note on Discrete Approximations of Continuous Distributions," 2006. Unpublished, University of Wisconsin–Madison.

_ , "Open Borders," *Review of Economic Dynamics*, April 2013, *16*, L1–L13.

_ **and James R. Walker**, "The effect of expected income on individual migration decisions," *Econometrica*, January 2011, *79* (1), 211–251.

McFadden, Daniel D, "Conditional logit analysis of qualitative choice behavior," in Paul Zarembka, ed., *Frontiers in Econometrics*, Academic Press, 1974.

McFadden, Daniel L., "Modelling the Choice of Residential Location," in F. Snickars A. Karlqvist, L. Lundqvist and J. Weibull, eds., *Spatial Interaction Theory and Planning Models*, North Holland: Amsterdam, 1978, chapter 3, pp. 75–96.

Rust, John P., "Structural Estimation of Markov Decision Processes," in R.F. Engle and D.L. McFadden, eds., *Handbook of Econometrics*, Vol. 4, Amsterdam, New York and Oxford: Elsevier Science, North-Holland, 1994, pp. 3081–3143.