SPATIAL VARIATION IN HIGHER EDUCATION FINANCING AND THE SUPPLY OF COLLEGE GRADUATES

JOHN KENNAN
UNIVERSITY OF WISCONSIN-MADISON AND NBER

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. In light of this finding, it is natural to analyze the quantitative implications of a uniform national tuition policy, and the model is well-suited for such an analysis. The estimated model is used to simulate the effects of a national tuition policy that would eliminate all cross-State tuition differences. The model predicts that under such a policy, there would be a large increase in the proportion of college graduates in the labor force. Moreover, although this involves a substantial reduction in tuition revenue, there is an almost equally large offsetting increase in income tax revenue, such that the policy virtually pays for itself.

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?

Date: May, 2020.
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 Amanda Agan, Kartik Athreya, Gadi Barlevy, John Bound, Rebecca Diamond, Eric French, Chao Fu, Manolis Galenianos, Ahu Gemici, Grey Gordon, Joseph Han, Jim Heckman, Tom Holmes, John Bailey Jones, Lisa Kahn, Maurizio Mazzocco, Bob Miller, Alex Monge-Naranjo, David Mustard, Derek Neal, Luigi Pistaferri, Jack Porter, Mike Rothschild, Jean-Marc Robin, Richard Rogerson, Chris Taber, Sarah Turner, Jim Walker, Yoram Weiss, Danny Yagan and many seminar participants for helpful comments.

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 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 levels 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 (or the best value for money), 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.

In light of this finding, it is natural to analyze the quantitative implications of a uniform national tuition policy, and the model is well-suited for such an analysis. In particular, as Knight and Schiff (2019) have pointed out, a national tuition policy would eliminate the distortions associated with the large differences between resident and non-resident tuition levels. The estimated model is used to simulate the effects of a policy that sets tuition in all States at the actual California resident level. The results are quite striking. The model predicts that under such a policy, there would be a large increase in the proportion of college graduates in the labor force. Moreover, although

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

this involves a substantial reduction in tuition revenue, there is an almost equally large offsetting increase in income tax revenue, such that the policy virtually pays for itself.

2. Geographical Distribution of College Graduates

There are 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.[2] 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.[3]

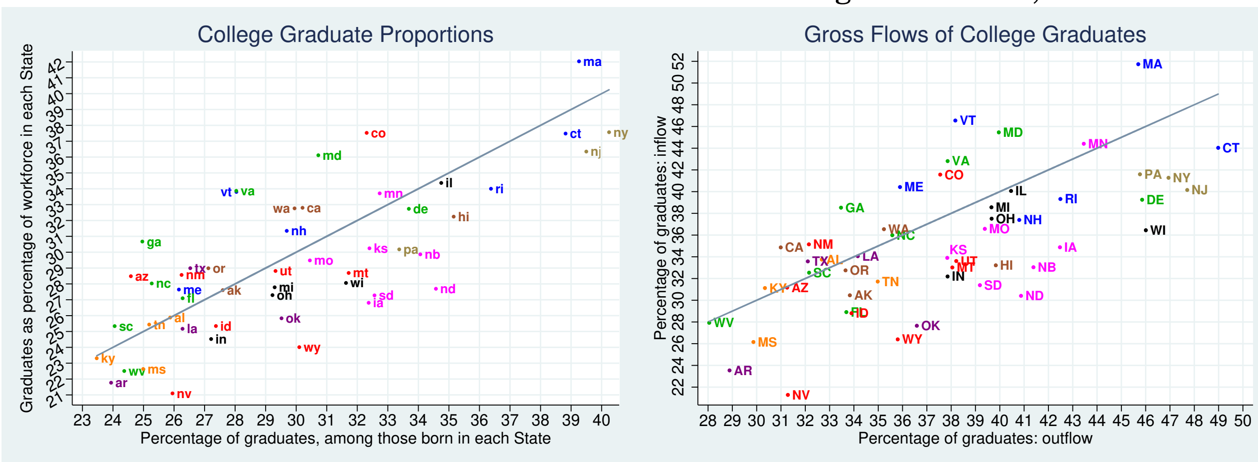
One might expect that the proportion of college graduates in the flow of in-migrants would be relatively high in States that have relatively few graduates in the native population, and similarly that the proportion of graduates in the flow of out-migrants would be high in States that have a high proportion of graduates in the native population. The right panel of Figure 2.1 shows that this is not the case.[4]

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

[3]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 the proportion working in their birth-State is above 65% for college graduates who were born in Texas or California, and it is below 25% for Alaska and Wyoming.
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, 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.

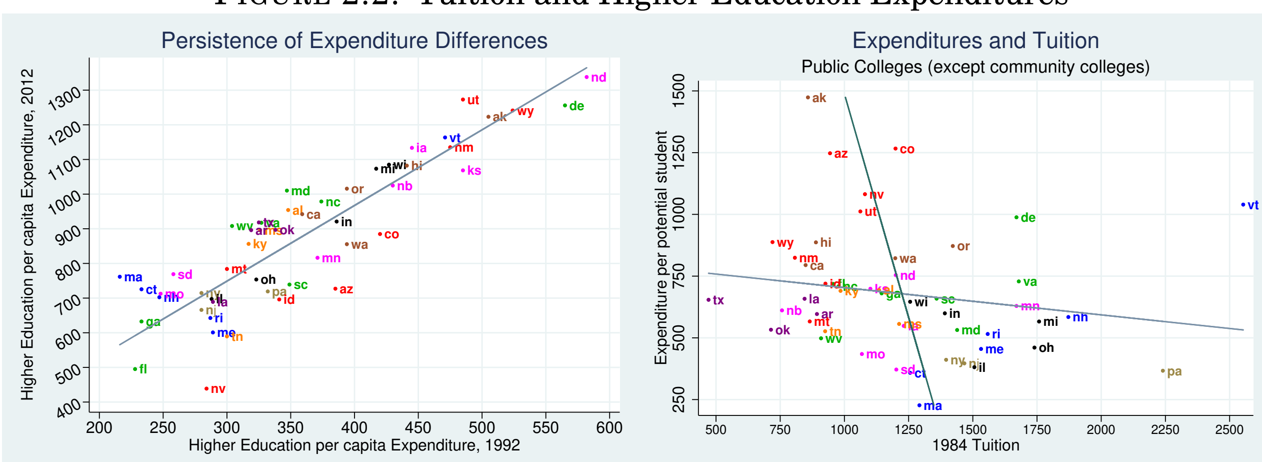
[4]One possible explanation for the differences in the proportion of college graduates across States is that there are similar differences in the proportion of college graduates in the parents' generation, and there is a strong relationship between the education levels of parents and children. Of course 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.

FIGURE 2.1. Birth and Work Locations of College Graduates, 2000



These figures show birth and work locations for labor force participants aged $25-50$ in the 2000 Census. On the left, the percentage of college graduates among those working in each State is plotted against the percentage among those born in that State. The figure on the right refers to gross flows: people born in one State who are working in a different State. In this figure, the percentage of graduates among the (domestic) immigrants to each State is plotted against the percentage of graduates among the emigrants.

FIGURE 2.2. Tuition and Higher Education Expenditures



The figure on the left uses data on State & Local Finances from the Census of Governments. Per capita State and local expenditures on higher education for each State in 2012 are plotted against the corresponding amounts for 1992, with a regression line superimposed. The figure on the right uses data from IPEDS (nces.ed.gov/ipeds). The y-variable is calculated by summing instructional expenditures over all public colleges in each State, and dividing by the number of high school graduates born in that State who were in the age range $22-36$ in the 1990 Census. The $x-$variable is an enrollment-weighted average of sticker prices in public colleges in each State in 1984. Two regression lines are superimposed (y on x and x on y).

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 perceived 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.[7]

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%.[8]

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

[6] Here 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] Ishimaru (2020) analyzes the importance of birth location for lifetime incomes, showing that spatial differences in college access and in local labor market conditions both have substantial effects on college graduation rates and earnings.

[8] 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

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 variables that are correlated with tuition levels.[9] 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. 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."

---

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.

[9]Kane (2006) gives the example of California spending a lot on community colleges while also having low tuition. Deming and Walters (2018) use tuition caps and cuts in State appropriations to estimate the effects of changes in expenditures and tuition levels. For example, there might be budget cuts in two States, with a tuition cap in one but not the other. Colleges can respond by cutting expenditure or increasing tuition. If some colleges are not allowed to increase tuition, the response by these colleges can be used to isolate the effect of expenditure changes. Then for unconstrained colleges, the expenditure effect is known, and any additional effect is attributed to tuition. A measure of previous reliance on State appropriations is used to instrument for budget cuts. Using this method, Deming and Walters (2018) find little evidence that tuition affects enrollment, but the contrast between this finding and the "consensus" view might well be explained by the use of a roundabout procedure that yields a relatively clean but quite noisy measure of exogenous variations in tuition levels.

[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.

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 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 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.

4.1. **Enrollment Decisions.** In simple models of higher education choices, high school graduates choose whether to forgo 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 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, or enroll in two-year colleges, with the possibility of subsequently transferring to a four-year college.[11] Accordingly, the model analyzed here treats college choices as a sequence of decisions on whether to enroll in one of several types of college, with uncertainty about whether enrollment will lead to graduation.

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 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 (one-year) period, there is a choice of whether to enroll in college, with the understanding that transitions from one schooling level to the next are stochastic. There are four college types: community colleges, other public colleges and universities, and private colleges at two quality levels.[12] The college types differ with respect to tuition, State subsidies, financial aid, transition 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.[13]

4.2. **Enrollment and Migration.** 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 net lifetime income. Let x be the state vector (which includes the stock of human capital, 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,[14] the utility flow is specified as $u(x, a) + \zeta_a$, where ζ_a is a random variable that is assumed to be i.i.d 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.

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

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

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

[14]See Rust (1994).

The decision problem can be written in recursive form as

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

where

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

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.[15] 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.

4.3. **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 location choices), and n columns (for enrollment choices).[16] 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, and $\kappa > 0$ is a parameter reflecting the relative predictability of migration and enrollment choices. Then if row i has been chosen, the column choice is determined by

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

[15]A simple derivation of this result can be found in Kennan (2008)

[16]Thus the ij array element refers to a person who chooses to live in location i, and who enrolls in a college of type j, with the understanding that if $j = 0$, this person is working rather than attending college.

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

$$i = \arg\max_{s} (\bar{v}_s + \zeta_s)$$

The probability of choosing row i is

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

and the expected value of the whole array is

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

The choice probabilities are then given by

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

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 $F(y)$ given by

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

subject to the restriction $0 < \kappa \leq 1$ (which ensures that the density function is non-negative).[17] 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 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).

4.4. **State Variables and Flow Payoffs.** Let $\ell = (\ell^0, \ell^1)$ denote the current and previous location, let ω be a vector recording wage information at these locations, and let (ξ, ξ^0) denote current and lagged enrollment choices (with the convention that $\xi = 0$ means non-enrollment, and otherwise ξ represents the college type).[18] The state vector x consists of ℓ, ξ^0, ω, current education level (e) and age (g).[19]

The deterministic part of the flow payoff is specified as

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

Here the first term refers to psychic benefits associated with different educational attainment levels. The second term refers to after-tax wage income in the current location (ignoring random components that do not affect enrollment and migration choices, as discussed in Section 4.5 below). The wage depends on age, schooling and ability, and $\mathbb{T}(\ell^0)$ is the tax rate in the current location. Income while in college is assumed to be zero (the indicator notation $\mathbb{I}(\xi = 0)$ means no college enrollment). Income is augmented by the amenity variable $Y(\ell^0)$. The parameter α^H is a premium reflecting a preference for the home location. The cost of attending a college of type ξ in location ℓ for a person with home location h is denoted by $C_h(\ell, \xi)$, with $C_h(\ell, 0) = 0$. The cost of moving to location j is represented by $\Delta(x, j)$. These college costs and moving costs are specified in Sections 4.6 and 4.7.

---

[17]See Börsch-Supan (1990). Note that the sequential choice interpretation relaxes the requirement that κ must be less than unity.

[18]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.

[19]There are also state variables that are fixed for each individual: home location, ability, mother's and father's education, and family income; in the individual's dynamic programming problem these are treated as parameters.

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

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

where e is schooling level, μ_j is the mean wage in location j (for each schooling level),[20] υ is a permanent location match effect, $G(e, b, g)$ represents the age-earnings profile for schooling level e and ability b, η is an individual effect that affects wages in the same way across all schooling and ability levels and across 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, and it is assumed that $\upsilon_{ij}(e_i)$ is realized only after moving to j.

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 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 $(\theta_e, y_e^*, c_e, \text{ and } g_e^*)$.[21]

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 education level. The individual effect (η) and the transient effect (ε) 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.[22]

4.5.1. *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 finite support, and compute continuation values at the support points. For given support points, the best discrete approximation $\hat{F}$ for any distribution F assigns

[20] Note that wages depend on schooling level without regard to the path by which this level was reached; in particular, there is no distinction between graduates of public and private colleges. The main reason for this is that although the NLSY data include information about college type, these data are not rich enough to generate reasonable estimates of wage dispersion across all States, and the Census data that are used for this purpose do not contain information about college type.

[21] 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.

[22] In particular, η is not a state variable in the dynamic programming problem. Note that the dispersion of the transient component may depend on the education level, but this has no relevance for the dynamic programming problem, since the objective is to maximize expected income.

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.[23] 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 individual 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. Accordingly, 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.

4.6. **College Costs.** Aside from expected income and the psychic benefits associated with educational attainment, 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 and on age (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, $E(\ell, \xi)$, financial aid (scholarships), $s(\ell, \xi)$, and parents' education and family income. Let d_m and d_f be

[23] See Kennan (2006)

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

$$\begin{aligned} C(x,\xi) &= \delta_0(\xi) + \delta_1(\xi)\,\tau(\ell,\xi) + \delta_2(\xi)\,E(\ell,\xi) + \delta_3 b + \delta_4 b\mathbb{I}(\xi \in \Xi) + \delta_5 d_m + \delta_6 d_f + \delta_7 y_f \\ &\quad + \delta_8(g - g_0) + (\delta_9(1-b)\,y_f + \delta_{10} b(1-y_f) + \delta_{11} b y_f + \delta_{12}(1-b)(1-y_f))\,s(\ell,\xi) \\ &\quad + \delta_{13}\mathbb{I}(e=1) + \delta_{14}\mathbb{I}\left(\xi = \xi^0\right) \end{aligned}$$

where tuition is given by

$$\tau(\ell,\xi) = \mathbb{I}(\ell = h)\,\tau_r(\ell,\xi) + \mathbb{I}(\ell \neq h)\,\tau_n(\ell,\xi)$$

(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 change as students advance (δ_{13}). [24] 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 exactly how this effect operates. The effect of scholarships is measured separately for each of the four ability and family income type combinations. 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, but this is more relevant for students who are more likely to be eligible for scholarships. Finally, persistence in enrollment choices is captured by the parameter δ_{14}.

4.7. **Moving Costs.** Moving costs are specified as in Kennan and Walker (2011).[25] 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

[24] 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 ability 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.

[25] The specification in Kennan and Walker (2011) used the convention that the anticipated sign of each coefficient should be positive. That convention is not used in this paper. Note also that there is no distinction between moves associated with out-of-State college enrollments and moves to a different labor market – the moving cost is the same in both cases.

cost is specified as

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

Thus the moving cost varies with education (γ_{0e}). The observed migration rate is much higher for college graduates than for high school graduates, and the model can account for this through differences in either the potential income gains or the cost of moving. The moving cost is an affine function of distance (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 area); a move to a previous location may also be less costly (γ_3). In addition, the cost of moving is allowed to depend on age. Finally, it may be cheaper to move to a large location, as measured by population size n_j.

4.8. **Transition Probabilities.** The state vector can be written as $x = (\tilde{x}, g)$, with $\tilde{x} = (e, \ell^0, \ell^1, \xi^0, x_v^0, x_v^1)$, where x_v^0 and x_v^1 index the realizations of the location match component of wages in the current and previous locations. Let $q_b(e, \xi, \xi^0)$ denote the probability of advancing from education level e to $e+1$, for someone with ability b who is not a college graduate, and who is enrolled in a college of type ξ, with $q_b(e, 0, \xi^0) = q_b(2, \xi, \xi^0) = 0$. The transition probabilities are as follows[26]

$$p(x' \mid x, a) = \begin{cases} q_b(e,\xi,\xi^0) & \text{if} \quad j = \ell^0, \quad \tilde{x}' = (e+1, \ell^0, \ell^1, \xi, x_v^0, x_v^1), \\ 1 - q_b(e,\xi,\xi^0) & \text{if} \quad j = \ell^0, \quad \tilde{x}' = (e \quad, \ell^0, \ell^1, \xi, x_v^0, x_v^1), \\ q_b(e,\xi,\xi^0) & \text{if} \quad j = \ell^1, \quad \tilde{x}' = (e+1, \ell^1, \ell^0, \xi, x_v^1, x_v^0), \\ 1 - q_b(e,\xi,\xi^0) & \text{if} \quad j = \ell^1, \quad \tilde{x}' = (e \quad, \ell^1, \ell^0, \xi, x_v^1, x_v^0), \\ \frac{q_b(e,\xi,\xi^0)}{R} & \text{if} \quad j \notin \{\ell^0, \ell^1\}, \quad \tilde{x}' = (e+1, j, \ell^0, \xi, x_v, x_v^0), \quad 1 \le x_v \le R \\ \frac{1 - q_b(e,\xi,\xi^0)}{R} & \text{if} \quad j \notin \{\ell^0, \ell^1\}, \quad \tilde{x}' = (e \quad, j, \ell^0, \xi, x_v, x_v^0), \quad 1 \le x_v \le R \\ 0 & \text{otherwise} \end{cases}$$

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

[26]There are six cases: either there is no move (the first two lines), or a move to a previous location (the next two), or a move to a new location (in which case there is a draw from the match component distribution). In each case these is also of course an age increment $(g' = g + 1)$.

through 1994 is used here, but wage information is available (biennially) through 2015, and this is used in order to obtain better estimates of the lifetime wage profile.

In order to analyze 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 12,893 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		
Community Colleges	469	17%
Other Public Colleges	1,499	53%
Private, lower tier	136	5%
Private, upper tier	736	26%
Subtotal	2,840	
Average years enrolled	3.7	
Not enrolled	10,023	
Total (person-years)	12,863	
Ever enrolled in college		
No	523	41%
In-State only	566	44%
Out-of-State only	98	8%
Both	94	7%
Total (persons)	1,281	

This table gives college enrollment counts for white non-Hispanic males in the core NLSY79 sample with at least 12 years of schooling, starting at age 19, and observed over the period 1979–1994. The classification of private colleges is based on the Carnegie classification (1994). The lower tier has four college types: Baccalaureate Colleges II, Associate of Arts Colleges, Schools of Business and Management and Tribal Colleges; all other private colleges are classified as upper-tier.

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).[27] The tax rate $\mathbb{T}(\ell^0)$ is measured as the sum of federal and State average tax rates for 1984, as calculated by the NBER TAXSIM model.[28]

[27] 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.
[28] See http://users.nber.org/~taxsim/allyup/.

5.1.1. *Tuition and Subsidies.* In the model, each State has one representative college of each type,[29] and all of these colleges are available choices for everyone.[30] Tuition rates were estimated by computing enrollment-weighted averages of real "sticker prices" for each college type, using averaged IPEDS data for 1980 and 1984. Students are assumed to pay tuition at the resident rate in their home State, while others pay the non-resident rate, allowing for a few reciprocity agreements.[31] The home State is defined as the State in which the individual finished high school.

Subsidy measures were constructed by adding federal, State and local appropriations and grants over all public colleges in the IPEDS files, 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.[32] 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

[29]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 model.

[30]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.

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

[32]These data were obtained from the IPEDS finance files for 1980 and 1984 (e.g 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). The IPEDS files for $1981-83$ are not available.

TABLE 2. **College Costs**

	Mean	S.D.
Tuition	($1983)	
Community Colleges, Resident	617	252
Community Colleges, Nonresident	1,708	658
Other Public, Resident	1,116	366
Other Public, Nonresident	2,902	847
Private, lower tier	3,507	846
Private, upper tier	4,705	1,571
Expenditure (per potential student)		
Community Colleges	103	89
Other Public Colleges	644	233
Private, lower tier	48	45
Private, upper tier	204	214
Financial Aid (per potential student)		
Community Colleges	8.3	6.2
Other Public Colleges	39.1	21.3
Private, lower tier	9.7	8.6
Private, upper tier	27.0	26.6

The unit of observation in this table is a State. The tuition levels are enrollment-weighted averages of real "sticker prices" within each State for the years 1980 and 1984. The expenditure and financial aid variables are totals over all colleges of each type within each State, divided by the number of high school graduates aged 22-36.

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, by ability level, where the ability measure is an indicator of whether the AFQT percentile score is above 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 of college-educated parents are much more likely to have high AFQT scores.

5.2. **Parameter Estimates.** Maximum likelihood estimates of the model parameters are shown in Tables 4, 5 and 6. The parameters of the wage process were estimated separately, using the most recent data (including the biennial interviews); these parameters were treated as known when estimating the parameters governing college choice and migration decisions.

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 **84.8%**	33 7.5%	34 7.7%	442 62.3%
High Ability	128 47.8%	56 20.9%	84 31.3%	268 37.7%
Total	503 70.8%	89 12.5%	118 16.6%	710
Both Parents went to College				
Low Ability	41 51.9%	19 24.1%	19 24.1%	79 29.7%
High Ability	24 12.8%	44 23.5%	119 **63.6%**	187 70.3%
Total	65 24.4%	63 23.7%	138 51.9%	266

Most of the parents of the $1,281$ men in the analysis sample did not go to college. In 710 cases, neither parent went to college, and the top panel shows educational outcomes for the sons of these parents, classified according to whether the AFQT score was above or below the median. In 266 cases, both parents went to college, and the bottom panel shows the outcomes for these cases.

TABLE 4. **Wage Parameters**

$$w_{ijt} = \mu_j(e_i) + \upsilon_{ij}(e_i) + \varepsilon_{ijt}(e) + \eta_i + \begin{cases} \theta_e b + y_e^* - c_e (g - g_e^*)^2 & g \le g_e^* \\ \theta_e b + y_e^* & g \ge g_e^* \end{cases}$$

	High School $\hat{\theta}$	High School $\hat{\sigma}_\theta$	Some College $\hat{\theta}$	Some College $\hat{\sigma}_\theta$	College $\hat{\theta}$	College $\hat{\sigma}_\theta$
Peak Wage (y_e^*)	1.80	0.014	2.36	0.040	2.35	0.040
Age at Peak (g_e^*)	36.88	0.366	43.93	1.311	49.36	0.692
Curvature (c_e)	-1.34	0.064	-1.00	0.138	-1.23	0.068
AFQT (θ_e)	0.17	0.022	0.20	0.044	0.16	0.040
Location match	0.45	0.008	0.95	0.021	0.92	0.016
Transient s.d. 1	0.60	0.002	0.83	0.007	0.79	0.005
Transient s.d. 2	2.70	0.009	3.93	0.055	4.44	0.028
$\bar{\mu}(e_i)$, s.d.	0.80	0.070	0.86	0.065	1.66	0.070

	All Education Levels $\hat{\theta}$	All Education Levels $\hat{\sigma}_\theta$
Individual Effect (η_i)	1.206	0.011

Wages are in levels at 1983 prices, in $\$10,000$ units. The individual effect η is uniformly and symmetrically distributed around zero, with three points of support; the estimated value of the positive support point is $\$12,060$ (regardless of education level). The transient component ε is drawn from a uniform mixture of two zero-mean normal distributions, with standard deviations that vary across education levels, ranging from $\$6,000$ to $\$44,400$. The $\bar{\mu}(e_i)$ row shows the (unweighted) average and standard deviation of the State wage levels. Thus for example the peak high school wage is $(0.80 + 1.80) \times \$10,000 = \$26,000$, on average across States, with a standard deviation of $\$700$.

The estimated wage function is increasing and concave in age (as usual), with a peak at age 37 for high school graduates, and at later ages for college-educated people. The direct effect of the (binary) AFQT score is relatively weak (conditional on the education level). Ability is of course strongly correlated with earnings, but the estimated earnings process attributes this largely to a strong relationship between ability and educational attainment. As was mentioned above, measured earnings are highly variable, both across people and from year to year for the same person, and this variability is reflected in the estimated dispersion parameters.

The estimated transition probabilities are shown in Table 5. Note that since the high school and some college levels span two years each, enrollment in two successive years is generally associated with much higher transition rates. For those enrolled in upper-tier private colleges for two consective years, the initial transition rates are higher for the high-ability group, but otherwise the relationship between ability and transition rates is generally weak.

TABLE 5. **Estimated College Transition Probabilities**

		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
Not enrolled in previous year					
Public	Lower-Tier	10.0%	0.0%	10.5%	0.0%
Public	Upper-Tier	10.3%	0.0%	12.5%	10.3%
Private	Lower-Tier	36.4%	0.0%	35.7%	25.0%
Private	Upper-Tier	14.6%	10.0%	4.8%	0.0%
Enrolled in previous year					
Public	Lower-Tier	43.4%	16.1%	51.6%	8.5%
Public	Upper-Tier	77.0%	37.1%	79.7%	36.5%
Private	Lower-Tier	57.1%	21.1%	81.0%	43.2%
Private	Upper-Tier	69.2%	33.3%	83.9%	37.7%

This table contains the estimated values of the parameters $q_b\left(e, \xi, \xi^0\right)$, which were specified in Section 4.8. For example, if someone in the lower half of the AFQT distribution who has completed 12 or 13 of schooling enrolls in an upper-tier public college, having been enrolled in some college in the previous year, there is a 77% chance that this person will have completed 14 or 15 years of schooling one year later. These estimates are the actual transition rates in the NLSY data used to estimate the model.

The main parameter estimates are shown in Table 6. 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.

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.

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

5.3. **Goodness of Fit.** Two goodness of fit measures are shown in Figure 5.1. The figures on the left show the actual and predicted numbers of college and high school graduates at age 28 (for States with at least 10 NLSY observations at this age), with regression lines and 45-degree lines superimposed.[33] Although the fit is imperfect, the model gives a reasonably good description of the cross-State variation in graduation rates. The figures on the right show the predicted and actual timing of college graduation or lack thereof. Again, the fit is reasonable, although the model misses the jump in

[33]Here high school graduates refer to people who have not advanced beyond the initial education level.

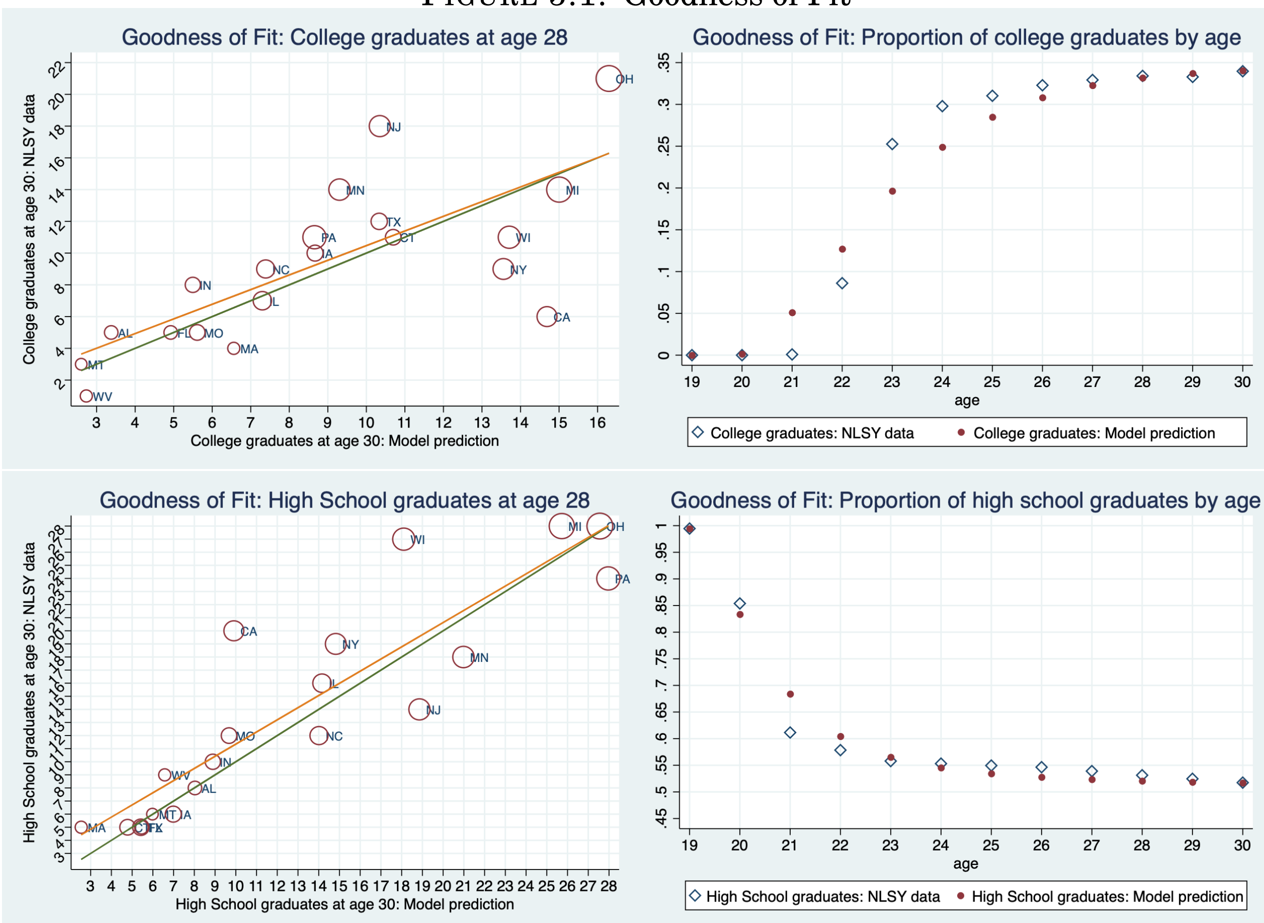
TABLE 6. **College Enrollment Choices and Migration**

Utility Parameters

		$\hat{\theta}$	$\hat{\sigma}_\theta$
A. Consumption value of education			
Some College	α_{01}	0.021	0.028
College Graduate	α_{02}	0.937	0.140
B. Flow Utility			
Income	α_1	0.109	0.010
Climate	α_2	0.011	0.004
Home Premium	α^H	0.175	0.009
C. Moving Cost			
High School	γ_{00}	4.665	0.255
Some College	γ_{01}	4.187	0.270
College Graduate	γ_{02}	4.175	0.286
Distance	γ_1	0.310	0.061
Adjacent Location	γ_2	-0.885	0.080
Previous Location	γ_3	-2.247	0.114
Age effect	γ_4	0.080	0.010
Population	γ_5	-0.853	0.064
D. College Cost			
Disutility, college: Pub lo	$\delta_0\,(1)$	2.196	0.169
Disutility, college: Pub hi	$\delta_0\,(2)$	1.999	0.199
Disutility, college: Pvt lo	$\delta_0\,(3)$	4.179	0.394
Disutility, college: Pvt hi	$\delta_0\,(4)$	3.003	0.269
Tuition: Public lo	$\delta_1\,(1)$	1.257	0.419
Tuition: Public hi	$\delta_1\,(2)$	2.166	0.200
Tuition: Private lo	$\delta_1\,(3)$	-1.647	0.588
Tuition: Private hi	$\delta_1\,(4)$	-0.103	0.259
Spend/Student: Pub lo	$\delta_2\,(1)$	-23.047	2.239
Spend/Student: Pub hi	$\delta_2\,(2)$	0.336	0.909
Spend/Student: Pvt lo	$\delta_2\,(3)$	-21.820	12.753
Spend/Student: Pvt hi	$\delta_2\,(4)$	-7.353	1.480
Ability effect on cost	δ_3	0.101	0.088
Ability×upper tier	δ_4	-0.376	0.116
Mother's education	δ_5	-0.071	0.027
Father's education	δ_6	-0.161	0.028
Family Income	δ_7	-0.040	0.032
Age	δ_8	0.105	0.008
Financial Aid (lo ab, hi inc)	δ_9	-10.780	9.719
Aid (hi ability lo income)	δ_{10}	-35.736	9.557
Aid (hi ability hi income)	δ_{11}	-53.106	9.461
Aid (lo ability lo income)	δ_{12}	11.883	13.611
cost diff, upper ed level	δ_{13}	-0.095	0.297
enrollment persistence	δ_{14}	-0.696	0.252
E. Payoff shocks			
Enroll/migrate shocks	κ	0.586	0.036

college graduation rates around age 23, as might be expected given that the model does not allow for heterogeneous transition probabilities.[34]

FIGURE 5.1. Goodness of Fit



The figures on the left plot the actual numbers of college and high school graduates at age 28 in the NLSY data against the numbers predicted by the model (for each State with at least 10 NLSY observations at this age); a regression line and a 45-degree line are superimposed. The figures on the right plot actual and predicted rates by age, using data for all States.

6. EFFECTS OF CHANGES IN TUITION LEVELS AND STATE EXPENDITURES

The results in Table 6 indicate that college enrollment decisions are affected by tuition and expenditure and financial aid levels, while expected income differences affect both

[34]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.

enrollment and migration decisions. The question then is whether changes in State policies regarding tuition and expenditures 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 Illinois 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 Illinois, 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 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 applied to the parental education data from the 1970 Census.

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.[35] 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.[36] 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

[35]Similarly, factor price equalization implies that immigration does not affect wages; see Kennan (2013)
[36]See Katz and Murphy (1992) and the many thousands of references thereto (scholar.google.com/citations?user=MAqEdFMAAAAJ).

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 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 in the data, and then doing the same thing for the new values of the policy variables, and comparing the population distributions.[37]

Some illustrative results are shown in Table 7, taking Illinois as an example; results for 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 of policy changes on tuition revenue, and on tax revenue (computed as present values over the lifecycle).

[37]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 associated with unlikely choices).

TABLE 7. **Effects of Policy Changes: (Illinois)**

		Population at Age 36						Revenue	
		Current Location			**Home Location**				
		Graduates	Some College	High School	Graduates	Some College	High School	Taxes	Tuition
Tuition Changes (20%)									
Decrease	Illinois	10.6%	4.1%	-5.5%	12.6%	4.2%	-5.9%	0.14%	-7.3%
	elsewhere	0.20%	0.02%	-0.07%	0.08%	-0.002%	-0.03%	0.02%	0.002%
Increase	Illinois	-9.3%	-4.1%	5.0%	-11.2%	-4.3%	5.4%	-0.07%	4.0%
	elsewhere	-0.17%	-0.03%	0.05%	-0.05%	0.002%	0.021%	-0.02%	-0.005%
Expenditure Changes (20%)									
All public colleges									
Increase	Illinois	1.65%	2.90%	-1.45%	2.07%	3.15%	-1.56%	-0.02%	2.50%
	elsewhere	0.023%	0.03%	-0.01%	0.002%	0.006%	-0.002%	0.005%	0.01%
Decrease	Illinois	-1.47%	-2.62%	1.30%	-1.85%	-2.84%	1.40%	0.02%	-2.22%
	elsewhere	-0.03%	-0.03%	0.01%	-0.002%	-0.006%	0.002%	-0.004%	-0.01%
Community Colleges									
Increase	Illinois	2.04%	3.04%	-1.66%	2.56%	3.30%	-1.78%	-0.02%	3.02%
	elsewhere	0.03%	0.03%	-0.01%	0.002%	0.003%	-0.002%	0.006%	0.01%
Decrease	Illinois	-1.86%	-2.76%	1.51%	-2.33%	-2.99%	1.62%	0.02%	-2.74%
	elsewhere	-0.03%	-0.03%	0.01%	-0.003%	-0.005%	0.002%	-0.005%	-0.01%

The estimated model is used to simulate decisions made by $300,000$ people of each type, distributed equally over the three wage bins, with each person starting in the home location at age 19. There are 800 types, classified by home location, and by four binary characteristics (ability, family income, and mother's and father's education). The resulting distribution over residence locations and education levels at age 36 is tabulated, using type weights corresponding to the distribution of types in the 1970 Census. This is done using the actual tuition and expenditure levels in the data, and re-done using alternative values of these policy variables, and the differences in the outcomes are shown in the Table. For example, when the tuition level (in public colleges) in Illinois is increased by 20% (with everything else held constant), the number of college graduates working in Illinois at age 36 falls by 9.3%, when compared with the corresponding number in the baseline simulation (using the actual tuition level). And among men whose home location was Illinois (regardless of where they were living at age 36), the number of college graduates falls by 11.2%.
The last two columns use the simulated data for the baseline and alternative levels of the policy variables to compare the net effects of 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 $.95$.

Tables 8 and 9 show that 20% tuition reductions[38] or expenditure increases (for public colleges) generate substantial (and roughly symmetric) changes in the proportions of college-educated men.[39] The tuition effects are generally larger than the expenditure

[38]The effects of tuition increases are tabulated in Appendix A.2.

[39]These policy changes induce some substitution between public and private colleges (since tuition and expenditure levels in private colleges are held constant). This effect is fully accounted for in the simulation results shown in the table (and in subsequent tables).

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 6). The estimated effects vary considerably across States. There are basically two reasons for this: a 20% change in tuition or expenditures corresponds to different dollar amounts, depending on the initial level, and the baseline proportions vary considerably across States.[40]

TABLE 8. **Effects of Tuition Reductions**(20%)

	Tuition (resident)	Grad Base	Population Changes (age 36): Current Location			Home Location			Revenue Changes (present values): Taxes		Tuition	
			Grad	Some College	High School	Grad	Some College	High School				
	$	%	Δ%	Δ%	Δ%	Δ%	Δ%	Δ%	Δ%	$	Δ%	$
California	655	40.8	3.9	-0.1	-4.1	3.8	-0.3	-4.5	0.15	44	-13.2	-95
Florida	946	27.4	7.0	2.4	-4.1	8.8	2.4	-5.0	0.17	10	-10.7	-27
Illinois	1317	23.7	10.6	4.1	-5.5	12.6	4.2	-5.9	0.14	24	-7.3	-43
Massachusetts	1197	40.9	2.7	0.8	-2.7	2.8	0.9	-3.0	-0.10	-11	-10.6	-33
Michigan	1640	19.2	15.1	5.9	-6.2	17.6	6.2	-6.5	-0.10	-17	-4.1	-24
New Jersey	817	18.9	12.1	5.5	-4.6	15.5	5.9	-4.9	-0.24	-24	-5.6	-18
New York	1391	31.1	8.5	2.7	-5.9	9.2	2.7	-6.3	-0.01	-4	-9.4	-105
North Carolina	747	23.6	6.3	1.9	-3.0	7.0	1.9	-3.4	-0.11	-10	-10.0	-15
Ohio	1565	16.2	14.9	7.1	-5.0	18.5	7.6	-5.2	-0.09	-17	-3.6	-20
Pennsylvania	2096	30.0	14.6	7.2	-4.8	20.4	8.0	-5.0	-0.14	-27	0.6	4
Texas	496	26.7	4.7	1.9	-2.5	5.4	1.9	-2.7	0.27	40	-13.7	-44

See the note below Table 7.
The second column shows resident tuition levels for upper-tier public colleges.
The third column shows the proportion of college graduates in the baseline simulation.

It may seem surprising that when a State reduces tuition, tax revenue received by that State falls in several cases, even though the proportion of college graduates increases, and college graduates pay more taxes (for example, this effect is particularly strong for Michigan). The main reason for this is that college graduates are more mobile, and some people who graduated because of the tuition reduction move to other States because they are graduates, and the tax revenue is then collected elsewhere. And since there are more graduates, and not all of them leave, the proportion of graduates can rise even though the tax base shrinks, and the net effect may be a reduction in total tax revenue. There is also a tendency to move to States where tax rates are lower, but the magnitude of this effect is small. If tax payments are classified by who paid rather than by who got the money, the effect of a reduction in tuition is an increase in tax payments.

[40]For example, the initial tuition levels and the college graduate proportions in the baseline simulation are shown in the first two columns of Table 8; a regression of the changes in the college graduate proportions (shown in the third column) on these two variables gives an R^2 of 0.91.

TABLE 9. **Effects of Increases in Community College Expenditures**(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%	$
CA	2.2%	3.1%	-4.2%	2.3%	3.2%	-4.7%	-0.08%	-24	1.3%	9
FL	2.5%	3.5%	-2.4%	-17.7%	-5.5%	-8.1%	-0.06%	-4	4.2%	11
IL	1.5%	2.2%	-1.2%	1.9%	2.4%	-1.3%	-0.02%	-3	2.2%	13
MA	0.4%	1.1%	-0.7%	0.5%	1.2%	-0.8%	-0.06%	-7	1.5%	5
MI	1.7%	2.6%	-1.1%	2.1%	2.8%	-1.2%	-0.04%	-8	2.2%	13
NJ	1.6%	2.4%	-0.9%	2.2%	2.7%	-1.0%	-0.09%	-9	2.6%	8
NY	1.1%	1.8%	-1.1%	1.2%	1.9%	-1.2%	-0.05%	-15	2.6%	29
NC	1.7%	2.7%	-1.4%	2.1%	3.0%	-1.5%	-0.13%	-12	1.6%	2
OH	1.5%	2.2%	-0.8%	2.0%	2.4%	-0.8%	-0.04%	-7	2.5%	14
PA	1.2%	2.0%	-0.6%	1.8%	2.2%	-0.6%	-0.03%	-6	2.18%	12
TX	1.0%	1.5%	-0.9%	1.3%	1.7%	-0.9%	0.01%	2	1.4%	5

See the note below Table 7

6.2.1. *Policy Effects and Migration.* A striking result in Table 8 is that changes in the policy variables have substantial long-term effects on the educational composition of the local population, despite some leakage due to migration. 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 little scope for State policies that aim to increase the proportion of college graduates 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 policy changes 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. For example, if the increased flow of graduates was largely due to increased numbers of 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.[41] Clearly, the increases in the proportion of college graduates in a State's labor force shown in Table 8 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 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).

TABLE 10. **Effects of Resident and Nonresident Tuition Reductions**(20%)

	Population (age 36)		Revenue		Population (age 36)		Revenue	
	Resident		Taxes	Tuition	Nonresident		Taxes	Tuition
	Graduates	Stayers	Δ%		Graduates	Stayers	Δ%	
CA	3.5%	81.3%	-0.03%	-16.0%	0.34%	44.8%	1.89%	10.1%
FL	6.1%	82.9%	-0.18%	-13.5%	0.89%	44.4%	1.43%	4.2%
IL	10.1%	77.3%	-0.03%	-8.8%	0.53%	35.9%	1.45%	8.9%
MA	2.4%	72.0%	-0.26%	-13.3%	0.29%	30.5%	1.14%	1.5%
MI	14.6%	72.8%	-0.26%	-5.4%	0.50%	28.2%	1.37%	15.7%
NJ	11.3%	73.3%	-0.47%	-6.9%	0.73%	28.9%	1.32%	2.8%
NY	8.1%	75.3%	-0.17%	-10.2%	0.40%	32.2%	1.52%	-0.4%
NC	5.7%	73.3%	-0.34%	-13.9%	0.62%	28.7%	1.35%	9.9%
OH	14.3%	73.8%	-0.24%	-4.8%	0.57%	31.1%	1.30%	11.0%
PA	14.1%	73.2%	-0.26%	-0.1%	0.51%	30.8%	0.86%	7.3%
TX	4.1%	83.9%	0.07%	-15.7%	0.63%	51.4%	1.52%	-2.6%

The "Stayers" columns show the proportion of the national increase in graduates found in the State where the tuition rate changed. Tuition and tax revenues are given as percentages of the base revenues for residents and nonresidents respectively.

6.2.2. *National Tuition Rates.* Given the sizable effects of tuition changes shown in Table 8, and the dispersion in tuition levels depending on State of residence, it is natural to consider the effects of a uniform national tuition policy.[42] For example, the largest State is California, and tuition levels there 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

[41]It is actually not necessary to run additional simulations to measure these effects. The results in Table 8 involve separable responses by two different groups of people to two components of the policy change: residents of the home State respond only to the level of resident tuition, and nonresidents respond only to nonresident tuition. The effects shown in Table 10 were obtained by disaggregating the results in Table 8 over these two groups.

[42]See Knight and Schiff (2019) for an analysis of the distortions associated with differential resident and nonresident tuition levels.

TABLE 11. **Effects of Revenue-Neutral National Tuition Reductions**

	Population at Age 36							
	Current Location			Home Location				
	College Graduates	Some College	High School	College Graduates	Some College	High School	Taxes	Tuition
CA	10.6%	-2.2%	-7.7%	6.6%	-3.3%	-6.4%	478	-46
FL	32.9%	6.3%	-15.0%	25.8%	4.8%	-13.9%	164	-117
IL	48.5%	8.6%	-21.6%	51.5%	8.6%	-22.0%	220	-238
MA	10.9%	-0.3%	-12.2%	11.5%	0.2%	-11.7%	-138	-144
MI	70.7%	14.7%	-26.1%	78.7%	15.3%	-26.8%	-44	-252
NJ	56.6%	13.4%	-19.8%	68.0%	14.7%	-20.0%	-110	-129
NY	36.0%	5.3%	-23.2%	36.4%	5.3%	-23.6%	46	-557
NC	19.6%	0.4%	-7.6%	15.5%	-0.4%	-6.4%	35	-25
OH	76.0%	20.8%	-22.6%	89.0%	22.5%	-23.0%	-33	-229
PA	93.6%	25.9%	-26.3%	119.1%	28.6%	-26.9%	12	-210
TX	11.5%	-1.6%	-2.5%	3.8%	-2.5%	-1.0%	269	0
US	35.2%	6.7%	-16.4%	35.2%	6.7%	-16.4%	3,589	-3,518

See the note below Table 7. Tuition levels in all States are set at the California in-State levels, and an income tax surcharge (0.741% of the prior tax) is imposed to balance the budget at the national level.

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 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 35%.[43] 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 big increase in the number of college graduates.[44] Moreover, the national tuition policy virtually pays for itself: the tax increase required to offset the lost tuition revenue is negligible (a 0.07% surcharge yields a small surplus).[45]

[43]This large effect involves several simultaneous changes: the distinction between resident and nonresident tuition levels is eliminated, cross-State differences in tuition levels are eliminated, and there is a general reduction in tuition levels. These effects are considered one at a time in Appendix A.3.

[44]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.

[45]The average tax rate in the baseline simulation is 8.918%; the results in Table 11 were generated by multiplying the tax rate in each State by a factor of 1.000741, meaning that the average tax rate rises to 8.925%.

6.2.3. *General Equilibrium Effects.* The estimates in Table 11 assume that changes in the proportion of college graduates have no effect on wages. But the changes are large, and 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 (in case the factor price equalization theorem is not applicable for some reason).[46] Suppose the ratio of college graduates to high school graduates increases by 43% (which is about the right magnitude, as will be shown). 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 8.6%. Table 12 shows how the results in Table 11 change when it is assumed that the college wage falls by 4.3%, and the high school wage rises by 4.3% (assuming that the effect on the wage premium is symmetric, and that wages for the some college group can be ignored). The increase in the proportion of college graduates is indeed attenuated due to the change in relative supplies, but the substantive conclusion is much the same as before: setting uniform national tuition levels at the California resident rates, with a tiny income tax surcharge to offset the reduced tuition income, yields a 26% increase in the proportion of college graduates.

TABLE 12. **GE-Adjusted Revenue-Neutral National Tuition Reductions**

	Population at Age 36							
	Current Location			Home Location				
	College Graduates	Some College	High School	College Graduates	Some College	High School	Taxes	Tuition
CA	4.7%	-3.3%	-0.7%	1.1%	-4.2%	1.0%	394	-76
FL	23.6%	3.5%	-9.8%	16.6%	2.0%	-8.6%	157	-125
IL	37.5%	5.8%	-16.7%	40.5%	5.9%	-17.1%	189	-259
MA	6.4%	-1.5%	-7.4%	7.0%	-0.9%	-6.9%	-93	-150
MI	59.3%	12.0%	-22.1%	66.9%	12.6%	-22.7%	3	-271
NJ	44.6%	9.6%	-15.5%	54.3%	10.7%	-15.8%	-32	-141
NY	30.0%	7.4%	-11.9%	34.7%	5.9%	-10.5%	83	-17
NC	10.9%	-2.9%	-3.1%	6.5%	-3.9%	-1.9%	121	-33
OH	62.2%	16.8%	-18.6%	73.9%	18.4%	-19.1%	68	-251
PA	79.2%	22.1%	-22.4%	102.2%	24.5%	-23.1%	72	-232
TX	2.3%	-4.9%	2.0%	-4.9%	-5.9%	3.5%	240	-22
US	26.2%	4.0%	-12.0%	26.2%	4.0%	-12.0%	3,898	-3,894

See the note below Table 7. Tuition levels in all States are set at the California in-State levels, and an income tax surcharge (0.38% of the prior tax, increasing the average rate from 8.92% to 8.95%) is imposed to balance the budget at the national level.

[46] The model assumes that labor supply is inelastic; if this assumption were relaxed, equilibrium wages would also be affected by supply responses to changes in wages and tax rates.

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 non-Hispanic male high school graduates. The results suggest that more generous subsidies do have significant effects on college enrollments, 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 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.

Appendix (not for publication)

Appendix A.1. Estimation Details

The model was estimated by maximum likelihood, using a slightly modified Newton algorithm. Value functions were computed by backward induction, and numerical derivatives were used to estimate the gradient of the loglikelihood. The outer product of the gradient was used to estimate the Hessian of the loglikelihood, and standard errors of the parameter estimates were obtained from the Hessian in the usual way. Using an MPI-Fortran program, on a cluster with 320 CPUs, the likelihood maximum was typically reached in less than 24 hours once reasonable starting values for the parameters had been found.

The wage process was estimated in the same way, using an ordinary Fortran program on a single computer; this required only a few hours.

Appendix A.2. Tuition Effects: Increases and Decreases

Table 13 compares the effects of tuition reductions shown in Table 8 with the corresponding effect for increases in tuition. As was mentioned in the text, the effects on educational outcomes are roughly symmetric. But the revenue effects are more complicated. An increase in tuition generally yields an increase in tuition revenue, but the elasticity of this response decreases as the tuition level increases, and at the highest tuition level, the response is actually negative: enrollment falls so much that revenue decreases. This is illustrated in the left panel of Figure A.2.1. The tax revenue responses are illustrated in the right panel of Figure A.2.1. In several cases, there is a negative relationship between tuition and tax revenues, for the reasons given in the text.

Figure A.2.1. **Effects of Tuition Changes**

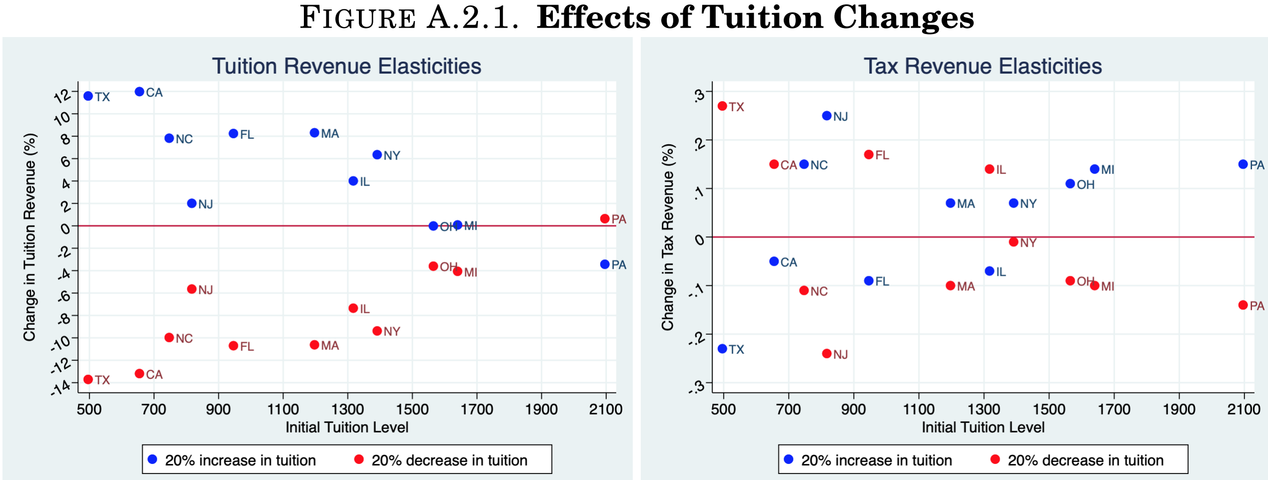


TABLE 13. **Effects of Tuition Changes**

			Population at Age 36						Revenue (present values)			
			Current Location			Home Location			Taxes		Tuition	
Tuition (resident)			College Graduates	Some College	High School	College Graduates	Some College	High School	Δ%	Δ$	Δ%	Δ$
CA	655	−20%	3.9%	-0.1%	-4.1%	3.8%	-0.3%	-4.5%	0.15%	44	-13.2%	-95
		+20%	-3.6%	0.1%	4.1%	-3.7%	0.2%	4.5%	-0.05%	-17	12.0%	86
FL	946	−20%	7.0%	2.4%	-4.1%	8.8%	2.4%	-5.0%	0.17%	10	-10.7%	-27
		+20%	-6.3%	-2.4%	3.9%	-8.2%	-2.6%	4.8%	-0.09%	-5	8.2%	21
IL	1,317	−20%	10.6%	4.1%	-5.5%	12.6%	4.2%	-5.9%	0.14%	24	-7.3%	-43
		+20%	-9.3%	-4.1%	5.0%	-11.2%	-4.3%	5.4%	-0.07%	-13	4.0%	23
MA	1,197	−20%	2.7%	0.8%	-2.7%	2.8%	0.9%	-3.0%	-0.10%	-11	-10.6%	-33
		+20%	-2.2%	-0.6%	2.1%	-2.2%	-0.6%	2.3%	0.07%	9	8.3%	26
MI	1,640	−20%	15.1%	5.9%	-6.2%	17.6%	6.2%	-6.5%	-0.10%	-17	-4.1%	-24
		+20%	-12.9%	-6.0%	5.6%	-15.3%	-6.4%	5.9%	0.14%	24	0.1%	0
NJ	817	−20%	12.1%	5.5%	-4.6%	15.5%	5.9%	-4.9%	-0.24%	-24	-5.6%	-18
		+20%	-10.3%	-5.3%	4.0%	-13.4%	-5.8%	4.4%	0.25%	25	2.0%	6
NY	1,312	−20%	8.5%	2.7%	-5.9%	9.2%	2.7%	-6.3%	-0.01%	-4	-9.4%	-105
		+20%	-7.5%	-2.7%	5.4%	-8.2%	-2.8%	5.7%	0.07%	22	6.3%	71
NC	747	−20%	6.3%	1.9%	-3.0%	7.0%	1.9%	-3.4%	-0.11%	-10	-10.0%	-15
		+20%	-5.8%	-1.9%	2.9%	-6.6%	-2.0%	3.2%	0.15%	13	7.8%	12
OH	1,565	−20%	14.9%	7.1%	-5.0%	18.5%	7.6%	-5.2%	-0.09%	-17	-3.6%	-20
		+20%	-12.3%	-6.7%	4.3%	-15.5%	-7.2%	4.5%	0.11%	22	0.0%	0
PA	2,096	−20%	14.6%	7.2%	-4.8%	20.4%	8.0%	-5.0%	-0.14%	-27	0.6%	4
		+20%	-11.1%	-6.3%	3.8%	-15.8%	-7.1%	4.0%	0.15%	29	-3.4%	-19
TX	496	−20%	4.7%	1.9%	-2.5%	5.4%	1.9%	-2.7%	0.27%	40	-13.7%	-44
		+20%	-4.4%	-1.9%	2.4%	-5.2%	-2.0%	2.6%	-0.23%	-34	11.6%	38

APPENDIX A.3. DECOMPOSITION OF TUITION EFFECTS

Table 12 shows that when tuition in every State is set at the California resident level the model generates a large increase in the proportion of college graduates. This involves the effects of several simultaneous changes: the distinction between resident and nonresident tuition levels is eliminated, cross-State differences in tuition levels are eliminated, and there is a general reduction in tuition levels. These effects are considered one at a time in Table 14. Since the point here is simply to provide some information about the sources of the large changes in graduation rates, the magnitudes of the separate effects are tabulated without adjusting for GE effects, or income tax rate changes.

With no adjustments for GE effects or tax rate changes, the overall effect of setting tuition at the California resident level is even larger than the effect shown in Table 12. A substantial part of the overall effect can be explained as a result of across-the-board reductions in tuition levels. This can be seen by considering the effects of a 20% tuition reduction in all States simultaneously (as opposed to the effects of a tuition reduction

TABLE 14. Decomposition of Tuition Effects

	CA Resident Tuition			20% Reduction			All Resident			National Tuition		
	Grad	Tax	Tuition	Grad	Tax	Tuition	Grad	Tax	Tuition	Grad	Tax	Tuition
CA	10.6%	446	-46	5.6%	95	-98	7.1%	356	-37	-14.2%	158	612
FL	32.9%	163	-117	11.0%	37	-28	10.7%	92	-4	-4.6%	8	15
IL	48.5%	199	-238	14.6%	68	-43	10.7%	0	-2	8.2%	41	-22
MA	10.9%	-149	-144	3.6%	-28	-36	1.5%	-74	-17	1.5%	18	-13
MI	70.7%	-60	-252	19.0%	-8	-25	10.2%	-106	11	25.7%	-1	-41
NJ	56.6%	-121	-129	16.5%	-17	-18	8.6%	-81	-4	12.5%	-30	-22
NY	36.0%	16	-557	10.8%	27	-108	5.9%	-170	-18	8.0%	1	-163
NC	19.5%	27	-25	9.4%	27	-15	8.3%	25	-4	-13.6%	84	57
OH	76.0%	-54	-229	19.8%	4	-19	12.6%	-93	10	23.7%	-6	-53
PA	93.5%	-10	-210	20.2%	6	5	13.0%	-93	10	39.7%	-53	-33
TX	11.5%	267	0	8.3%	109	-45	9.8%	166	-22	-23.7%	-118	155
U.S.	35.2%	3261	-3519	11.4%	991	-785	9.0%	1069	-127	0.2%	10	614

See the note below Table 7. In this table, there is no tax surcharge, and wages are not adjusted for GE effects. In the first panel, tuition levels in all States are set at the California in-State levels, as in Table 12. In the second panel, a 20% tuition reduction is applied in all States. In the third panel, tuition in all States is set at the in-State rate (for that State). In the last panel, all cross-State tuition differences are eliminated by setting tuition levels at the national average.

in a single State, as in Table 8). Next, simply eliminating the tuition surcharge paid by nonresidents generates a substantial increase in the proportion of college graduates. Finally, a uniform tuition policy that sets tuition at the national average level has a negligible overall effect, although there are big distributional effects (because the national policy means big increases in some States, and big reductions in other States).

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.

Deming, David J. and Christopher R. Walters, "The Impact of State Budget Cuts on U.S. Postsecondary Attainment," February 2018.

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.

Ishimaru, Shoya, "Geographic Mobility of Youth and Spatial Gaps in Local College and Labor Market Opportunities," 2020. sites.google.com/view/ishimaru/research.

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.

______, "Average Switching Costs in Dynamic Logit Models," 2008. 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.

Knight, Brian and Nathan Schiff, "The Out-of-State Tuition Distortion," *American Economic Journal: Economic Policy*, February 2019, *11* (1), 317–50.

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

______, "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.